Rochester Institute of Technology
RIT Scholar Works

Theses | Thesis/Dissertation Collections

11-21-2014

Perceptual Modeling and Reproduction of Gloss

Adria Fores Herranz

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation
Fores Herranz, Adria, "Perceptual Modeling and Reproduction of Gloss" (2014). Thesis. Rochester Institute of Technology. Accessed from

This Dissertation is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

Perceptual Modeling and Reproduction of Gloss

by

Adria Fores Herranz

B.S. Universitat de Girona, 2007
M.S. Universitat de Girona, 2009

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Program of Color Science
College of Science
Rochester Institute of Technology

November 21, 2014

Signature of the Author _____

Accepted by _____
Director, Ph.D. Degree Program Date

PROGRAM OF COLOR SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Adria Fores Herranz
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation required for the
Ph.D. degree in Color Science

Dr. Mark D. Fairchild, Dissertation Advisor

Dr. Reynold Bailey

Dr. Joe Geigel

Dr. Ingeborg Tastl

Date

Perceptual Modeling and Reproduction of Gloss

by

Adria Fores Herranz

Submitted to the
Program of Color Science
in partial fulfillment of the requirements
for the Doctor of Philosophy Degree
at the Rochester Institute of Technology

Abstract

The reproduction of gloss on displays is generally not based on perception and as a consequence does not guarantee the best visualization of a real material. The reproduction is composed of four different steps: measurement, modeling, rendering, and display. The minimum number of measurements required to approximate a real material is unknown. The error metrics used to approximate measurements with analytical BRDF models are not based on perception, and the best visual approximation is not always obtained. Finally, the gloss perception difference between real objects and objects seen on displays has not sufficiently been studied and might be influencing the observer judgement.

This thesis proposes a systematic, scalable, and perceptually based workflow to represent real materials on displays. First, the gloss perception difference between real objects and objects seen on displays was studied. Second, the perceptual performance of the error metrics currently in use was evaluated. Third, a projection into a perceptual gloss space was defined, enabling the computation of a perceptual gloss distance measure. Fourth, the uniformity of the gloss space was improved by defining a new gloss difference equation. Finally, a systematic, scalable, and perceptually based workflow was defined using cost-effective instruments.

Acknowledgements

I would like to express my gratitude to the following people who helped me during the course of my Ph.D.

First and foremost, I would like to thank my advisor Professor Mark D. Fairchild for his support and guidance during my studies. Professor Fairchild stimulated my interest in Color Science and gave me incredible support during the key moments of my graduate studies.

I would also like to express my gratitude to Dr. Ingeborg Tastl and John Recker from Hewlett Packard Laboratories, their direction and support both while interning at HP Labs and during my studies helped me immensely. I would also like to thank Hewlett Packard Laboratories for the support during my Ph.D.

I would like to thank Dr. James Ferwerda and Dr. Jinwei Gu for stimulating my curiosity at the intersection of perception and graphics and for advising me at the beginning of my Ph.D. I also would like to express my gratitude to my thesis committee members Dr. Joe Geigel and Dr. Reynold Bailey for their time and input.

Thanks also go to the Munsell Color Science Laboratory faculty, staff, and students. The experience would not have been the same without all of you, thanks!

Additionally, I would like to thank Françoise Viénot from the Centre de Recherche sur la Conservation des Collections in Paris, France, for allowing me to use the Eldim EZContrast and her assistance during the measurement process, and to Rhopoint Instruments for giving me access to the RAW data of the Rhopoint IQ, which enabled the last project of this thesis.

Finally, my deepest gratitude goes to my family. I would like to thank my wife Blanca and my family for their love and encouragement throughout the years.

To Blanca

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The reproduction of gloss on displays is generally not based on perception and as a consequence does not guarantee the best visualization of a real material. For example, the minimum number of measurements required to approximate a real material are unknown, the best visual approximations are not always obtained with currently used error metrics as they are not perceptually based, and the gloss perception in different media has not been sufficiently studied and might be influencing observer judgements.

The current situation of gloss is comparable in some regards to color decades ago when no metrics to determine the perceived difference between colors were available.

The visualization and communication of gloss could be greatly improved with a perceptually based reproduction workflow. Digital content creators like game and movie studios would be able to use the appearance of real materials in their digital models, while companies could provide a preview of the appearance of their products to customers. For example, customers could preview their content with different substrates or inks to aid the selection process and increase the confidence in the final product, and the same could be done for material selection in 3D printing applications.

Figure 1.1 shows the pipeline used to represent real materials on displays. First, the

material is measured. Second, the measured data is approximated with a mathematical model. Then, the synthetic image is rendered and finally displayed on a monitor.

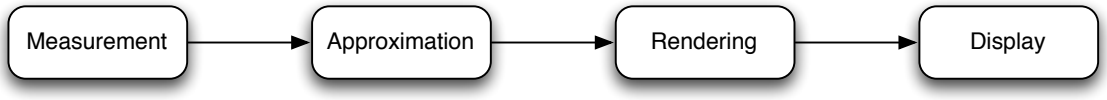


Figure 1.1: Pipeline to reproduce a real material on a display.

Gloss can physically be described by the Bidirectional Reflectance Distribution Function (BRDF), which describes the light-material interaction. This function is a 4-dimensional function that explains how the light is reflected at a given point of the surface for any incident and outgoing light direction. The measurement of the full 4-dimensional function is a challenging task, and only a subset of the function is commonly measured. As the most important perceptual features of gloss are not clearly defined, it's not clear that a specific measurement would lead to a good visual representation of the real material.

Analytical BRDF models are commonly used for rendering, as they present a compact representation and importance sampling can be used to speed-up the rendering process. For that reason, measured data is commonly approximated with analytical models to achieve those benefits while representing the measured material. However, the error metrics used in the optimization procedure do not have a perceptual basis and the obtained results do not always correspond to the best visual match.

Finally, rendering techniques are used to generate a synthetic image of the measured material, which is commonly displayed on a monitor. However, there have been no studies regarding the gloss perception between real materials and the representation of those materials seen on a monitor. Color appearance models, for example, describe the perception of colors depending on the viewing conditions and media used. Whether the different viewing conditions or media effect the gloss perception is currently not well understood.

This thesis focuses on five different projects leading to a better understanding and increased applicability of gloss perception:

- **Media dependent gloss perception:** The gloss matching performance of observers using real objects was first studied. Then, the same experiment was repeated with synthetic images. Finally, a cross-media matching experiment was performed, where the observers had to match a real material with synthetic representations. This project enabled the evaluation of the observer's matching performance in different media, and helped in the understanding of how well images approximate the gloss of real objects.

- **Perceptual evaluation of error metrics:** The different error metrics commonly used in the literature were perceptually evaluated. A paired comparison with reference study was performed, where the measured data was used to render the reference, and the observer was asked to select which of the two test images was closest to the reference. This project improved the understanding of which features are important in a metric to preserve the material appearance of approximations.

- **Gloss space projection for arbitrary BRDFs:** A projection for arbitrary BRDFs into the gloss space defined in Pellacini et al. [2000] using appearance ASTM standards was defined and validated in this project. By performing the projection of measured data and analytical models into the gloss space, the perceptual gloss difference between a real object and its representation in synthetic images can be evaluated. This projection can also be used to approximate a set of BRDF measurements without the need to perform non-linear optimization.

- **Space uniformity improvement:** In the previous project, the gloss space defined in Pellacini et al. [2000] was found to be non-uniform outside the region where the samples used to create the space are located. In this project the uniformity of the

space around additional gloss centers was evaluated, and an improved gloss difference equation that takes into account the space non-uniformity is presented.

- **Abridged gloss measurement:** Capturing the full BRDF is not scalable as it requires to store too much data, is time consuming, and because of the expensive nature of the devices required. A technique to measure the appearance of real objects using cost-effective instruments is presented in this project. This technique is based in splitting the material information to capture into two main attributes: color and gloss, and is evaluated by comparing its results with high accuracy measurements of a goniospectrophotometer.

In summary, the contributions from this thesis are the following:

- Evaluation and modeling of the media dependent gloss perception of real objects and synthetic images seen on displays.
- Perceptual evaluation of the error metrics currently used to approximate measured materials.
- The creation of a projection of arbitrary BRDFs into a gloss space, enabling to represent both, measured data and analytical models, and allowing the computation of the perceived difference between them.
- The definition of an improved gloss difference equation that takes the non-uniformity of the gloss space into account.
- A systematic, scalable, and perceptually based workflow using cost-effective instruments.

1.1 Thesis Overview

This thesis is based on five interrelated projects:

1. The study of gloss discrimination of observers on real objects and synthetic images is studied in Chapter 3. Matching experiments were performed with real objects, synthetic images, and in the cross-media scenario. This project enabled to understand how well images represent real objects' gloss and the gloss discrimination of observers in the different scenarios.

2. Metrics used to approximate BRDF measurements are not perceptually based, and for that reason a perceptual evaluation of commonly used BRDF metrics was performed in Chapter 4. This project helped to understand which metrics performed better from a perceptual point of view.

3. A projection for arbitrary BRDFs into the Pellacini et al. [2000] gloss space was defined in Chapter 5, allowing to access the space with analytical BRDF models and measured data, and to compute a perceptual distance measure.

4. A new gloss difference equation that takes into account the space non-uniformity of the gloss space defined in Pellacini et al. [2000] was defined in Chapter 6.

5. Finally, a technique to simplify the material appearance capture by using simpler instruments and the knowledge acquired in the previous projects was defined in Chapter 7, obtaining a cost-effective, systematic, and scalable measurement process.

Chapter 2

Literature Review

The interaction of a material with light is what defines its appearance. This topic has been greatly studied in the computer graphics community and in the appearance industry, the first community being more interested in representing real wold materials in synthetic images while the second has generally been more interested in quality control applications.

In this chapter the gloss perception from a perceptual standpoint will first be reviewed. Next, the physical description of the light-material interaction will be examined. Then, the measurement approaches that have been used in the computer graphics literature will be reviewed, and the efforts on approximating measured materials with analytical models will be described. Finally, the perceptual gloss spaces that have been created in the literature will be discussed.

2.1 Gloss Perception

Color, gloss, texture, and translucency are the main properties used to describe the material appearance of a given object [CIE, 2006]. Gloss and in some aspects also color are going to be the main focus of this thesis, without considering texture or translucency.

Hunter and Harold [1987] described seven features that relate to the perception of gloss:

Specular gloss This property models the specular reflection at different angles from the normal direction, commonly 20°, 30°, 45°, 60°, and 75°. Lower angles are used to compare high-specular materials and higher angles are used to compare low-specular materials.

Sheen This property models the specular reflection at grazing angles and it is defined at 85°.

Contrast gloss or luster Defines the difference between the highlight areas and its surrounding. This effect can clearly be seen in velvet cloth, which has distinct highlights and dark areas.

Absence-of-bloom gloss Also known by absence of haze, which is defined as the spread of the specular component of the reflected light from a glossy surface.

Distinctness-of-image gloss This property defines how clearly the surrounding environment is reflected on the surface of a material. For example, a mirror will have higher distinctness-of-image than a brushed metal as the mirror is going to reflect a sharp background, while the brushed metal will introduce some amount of blur to the reflected image of the background.

Surface-uniformity gloss This property defines how smooth a surface is, being able to perceive a non-uniform texture when the surface is rough.

Directionality This effect distinguishes the materials that are perceived equally if they are rotated over the ϕ angle from the ones that behave differently depending on the light source direction. This effect is normally described as isotropy and anisotropy, respectively.

The importance of gloss in the finishing of commercial products drove the creation of international standards concerning the measurements of some of those perceptual gloss attributes: Specular Gloss is defined in ISO 2813, ISO7668, ASTM D523, ASTM D2457, DIN 67530, and JIS 8741, Distictness-of-image gloss is defined in ASTM D5767, and Haze is defined in ASTM E430, and ISO 13803.

2.2 Bidirectional Reflectance Distribution Function

The Bidirectional Reflectance Distribution Function (BRDF), as described in Nicodemus et al. [1977], defines the light-material interaction for a material.

In order to decrease the dimensionality of the BRDF function, some effects like phosphorescence (variation over time), fluorescence (reemission of light at other wavelengths), surface non-uniformity, and subsurface scattering will not be studied in this thesis.

With those simplifications, the BRDF is a 4-dimensional function that describes how light is scattered by a surface and it is defined by the following equation:

$$f(\omega_i, \omega_o) = \frac{L(\omega_o)}{E(\omega_i)} \tag{2.1}$$

where E defines the irradiance due to the light source in the incoming direction defined by ω_i, and L defines the radiance of a surface in the outgoing direction ω_o. Directions are defined in spherical coordinates, where the angle from the normal direction (Z) is represented by θ, and the rotation along the XY plane is represented by ϕ. Only isotropic materials will be considered in this thesis, where the BRDF of a given material is independent of ϕ_i .

To correctly represent the light behavior a BRDF must fulfill the following laws:

Helmholtz reciprocity: The BRDF should be symmetric relative to the incoming and the outgoing direction of light. In other words, the same result must be obtained by

swapping the incoming and the outgoing directions:

$$f_r(\omega_i, \omega_o) = f_r(\omega_o, \omega_i) \tag{2.2}$$

Energy conservation: Materials must not add energy, except for light sources. In addition, the total amount of light reflected by a material can not exceed the total incoming light. This means that the ratio of reflection must be ≤ 1:

$$\int_{\Omega_o} f_r(\omega_i, \omega_o)(\omega_o \cdot n) d\omega_o \leq 1 \tag{2.3}$$

Using the reciprocity principle, one may also write:

$$\int_{\Omega_i} f_r(\omega_i, \omega_o)(-\omega_i \cdot n) d\omega_i \leq 1 \tag{2.4}$$

2.2.1 Measurement

The angular nature of the Bidirectional Reflectance Distribution Function makes BRDF measurements a challenging task. Goniometers make use of moving arms to place the light source and the sensor at the desired position. As the light source and sensor need to be repositioned for each set of angles (ω_i,ω_o), the measurement process becomes time consuming.

Because of this complex measurement task, multiple techniques have been developed to speed-up the measurement process. Normally, the efficiency gain comes at the expense of constraining the information being measured. Some techniques only measure in-plane information, where ϕ is restricted to 0° or 180°, while other techniques consider the material to be isotropic.

Both, commercial and lab-grade techniques exist to measure the BRDF of a given sample.

Next, an overview of lab-grade techniques that have been developed in the literature are presented.

An efficient goniometer design was introduced in Ward [1992]. The device was composed of a half-silvered hemisphere, a camera with a fisheye lens, and a moving light. The camera and the material sample were placed inside the half-silvered hemisphere, and a light source on a moving arm was placed outside the hemisphere. In each measurement, the light source was reflected off the sample and the reflection pattern was projected into the half-silvered hemisphere. The camera had a wide viewing angle using the fisheye lens, allowing to capture all the outgoing directions, projected in the hemisphere, at once.

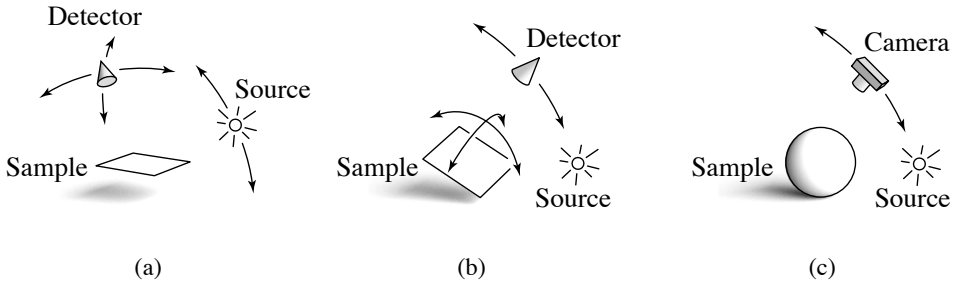


Figure 2.1: (a) Isotropic gonioreflectometer scheme, (b) Dana et al. [1999b] approach, and (c) Marschner et al. [1999] design. Image from Marschner et al. [1999].

Figure 2.1 shows various setups used for isotropic BRDF measurements. The goniometer seen in Figure 2.1(a) is composed of a moving light source and sensor. The design seen in Figure 2.1(b) has a fixed light source, a detector that can move in one direction and the ability to change the orientation of the sample. Finally, Figure 2.1(c) uses a spherical sample in combination with a digital camera to be able to acquire many samples in a single measurement.

The system seen in Figure 2.1(b) was described in Dana et al. [1999b], it involved mounting a planar material sample on a robot arm, which allowed to change the sample orientation. A sensor was moved to seven different positions, while collimated light was obtained using

a Fresnel lens. The CUReT BRDF database [Dana et al., 1999a] is available online and contains approximately 200 reflectance measurements for each of the 60 materials measured with this setup.

Figure 2.1(c) shows the design presented in Marschner et al. [1999]. In this case, a spherical sample and a camera are used instead of a planar sample to enable measuring multiple outgoing directions at once. In comparison with previous designs, a part from a faster measurement it also provides a much denser measurement sampling. Multiple images with the light source located at a different positions are taken to measure the three dimensional domain of the isotropic BRDF.

A similar system was used in Matusik et al. [2003] to create the publicly available MERL database, composed of more than 100 isotropic materials, including painted surfaces, fabrics, metals, and plastics. A sphere of a given material lit by a point light source was imaged with a camera for a dense set of incident directions. For each incident direction, a set of images with different exposures were merged to obtain an HDR image. This image-based method allowed to obtain high angular resolution measurements to the point of being able to render the measured data directly.

To overcome the limitation of only being able to measure isotropic materials with the above technique, a cylinder wrapped with strips of a material cut in different orientations was used in Ngan et al. [2005]. In this case, the different orientations of the strips are measured at once with a camera, thus allowing to have multiple in-plane measurements of an anisotropic material (see Figure 2.2).

2.2.2 Analytical BRDF models

Analytical BRDF models describe the light-material interaction with mathematical equations, allowing a fast evaluation and compact representation.

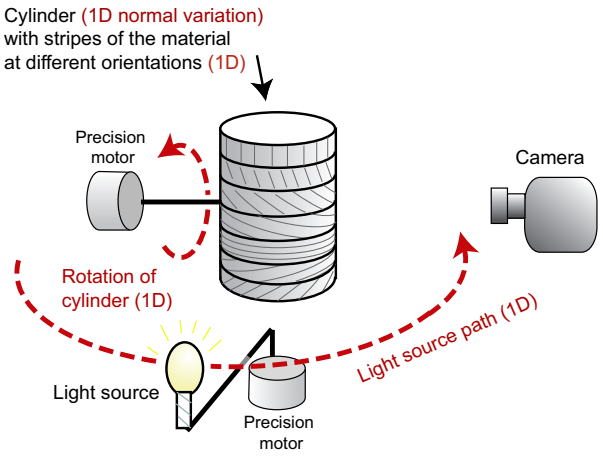


Figure 2.2: Acquisition setup for anisotropic materials. Image from Ngan et al. [2005].

Three different principles have been used to derive analytical BRDF models: Ideal material modeling, for which ideal diffuse and specular surfaces were modeled. Empirical modeling, for which ad-hoc models were created to either mimic the light-material interaction seen on real materials or to better approximate measured data. Physically based modeling, for which the definition of the underlying material microstructure was used to derive the model.

Next, the most common analytical BRDF models are described.

Lambertian Model

Lambertian, or "ideal diffuse", surfaces reflect light in all directions with the same radiance, having a constant $f_r(\omega_i, \omega_o)$. Using the conservation of energy principle, it can be shown that $f_r(\omega_i, \omega_o) \leq 1/\pi$.

In spite of being the model more widely used to represent diffuse materials in computer graphics, it can provide an inaccurate approximation of rough surfaces, as it doesn't model the retroflection effect seen when the viewing direction is closer to the incident light direction. The model presented in Oren and Nayar [1994] accounts for this effect and can improve the representation of those materials.

The opposite of a lambertian material would be an ideally specular material, in which all the incident light is reflected in the mirror direction.

Phong Model

The Phong BRDF model [Phong, 1975], also called cosine lobe, is based on the visual observation that shiny surfaces have specular circular highlights along the mirror reflection direction, and it tries to reproduce this behavior.

This model is defined by the following equation:

$$f_r(\omega_i, \omega_o) = \rho_s \frac{(R(\omega_o, n) \cdot \omega_i)^e}{(n \cdot \omega_i)} \tag{2.5}$$

where R is the reflection direction, n is the normalized surface normal, ρ_s defines the lobe magnitude, and e is the exponent that describes its sharpness. This model is not reciprocal and does not conserve energy.

Blinn-Phong Model

The Blinn-Phong BRDF model [Blinn, 1977] is a modification of the previous model. The model uses the halfway vector

$$h = \frac{\omega_i + \omega_o}{|\omega_i + \omega_o|} \tag{2.6}$$

which results in a more realistic modeling of the reflection behavior compared to the use of the viewing and reflection vector as in the original Phong model. The model equation is the following one:

$$f_r(\omega_i, \omega_o) = \rho_s \frac{(h \cdot n)^e}{(n \cdot \omega_i)} \tag{2.7}$$

This is the basic shading algorithm implemented by default in OpenGL and Direct3D's

fixed-function pipeline (before Direct3D 10), but it is neither reciprocal nor energy conserving.

Cook-Torrance Model

The Cook-Torrance BRDF model [Cook and Torrance, 1982] is a physically based model based on geometric optics principles, and models the surface microstructure by means of microfacets assuming that the scale of the surface roughness is large enough with respect to the wavelength of light. In the Cook-Torrance model, microfacets are assumed to be specular and the model is defined by:

$$f_r(\omega_i, \omega_o) = \frac{\rho_s}{\pi} \frac{DG}{(n \cdot \omega_i)(n \cdot \omega_o)} Fresnel(F_0, \omega_o, h) \tag{2.8}$$

where the Beckmann distribution is used to represent the normal distribution probability for the microfacets, D:

$$D = \frac{1}{\alpha^2 cos^4 \delta} e^{-[(tan\delta)/\alpha]^2}, \quad \delta = acos(n \cdot h) \tag{2.9}$$

α describes the surface roughness of the material, δ is the angle between the normal and the halfway vector, and G is the geometric attenuation term, which describes the masking and shadowing effects between the microfacets:

$$G = min\left(1, \frac{2(n \cdot h)(n \cdot \omega_o)}{(\omega_o \cdot h)}, \frac{2(n \cdot h)(n \cdot \omega_i)}{(\omega_o \cdot h)}\right) \tag{2.10}$$

The main features of this model is that it represents off-specular peaks and the increase in specular reflectance as the incident angle increases.

Commonly, the Fresnel term is approximated using the Schlick's approximation [Schlick,

1994], which depends on the parameter F_0:

$$Fresnel(F_0, \omega_o, h) = F_0 + (1 - F_0) \cdot (1 - (\omega_o \cdot h))^5 \tag{2.11}$$

Ward Model

An empirical anisotropic BRDF model was introduced in Ward [1992]. In this thesis, the isotropic version defined with the following equation will be used:

$$f_r(\omega_i, \omega_o) = \frac{\rho_s}{\sqrt{(n \cdot \omega_i)(n \cdot \omega_o)}} \cdot \frac{e^{-\tan^2 \delta / \alpha^2}}{4\pi\alpha^2} \tag{2.12}$$

where δ is angle between n and h, and α controls the width of the lobe.

Lafortune Model

The generalization of the cosine lobe was presented in Lafortune et al. [1997]. The new model is physically plausible, as it is energy conserving and reciprocal. The following equation is used to represent a Lafortune lobe:

$$f_r(u, v) = \rho_s[C_x u_x v_x + C_y u_y v_y + C_z u_z v_z]^m \tag{2.13}$$

where (u, v) are the incoming and outgoing directions in cartesian coordinates, C_x, C_y and C_z models the lobe behavior according to these directions, and m models the lobe's sharpness.

An isotropic behavior is obtained when $C_x = C_y$, otherwise the BRDF becomes anisotropic. Previously existing models can be obtained with this model. For example, the original cosine lobe model is obtained by choosing $-C_x = -C_y = C_z = \sqrt[n]{C_s}$, C_s being the normalization factor $(n + 2)/(2\pi)$ in order to preserve energy. The Lambertian model is obtained when

$n = C_x = C_y = 0$.

A linear combination of multiple lobes is also possible by using the following equation:

$$f_r(u,v) = \sum_{i}^{nLobes} [C_{x,i}u_x v_x + C_{y,i}u_y v_y + C_{z,i}u_z v_z]^{m_i} \tag{2.14}$$

When $-C_x = -C_y = C_z$, the lobe is centered on the reflection direction like the cosine lobe. If $C_{xy} > C_z$, the lobe goes under the mirror reflection direction, and for $C_{xy} < C_z$, the lobe goes over the mirror reflection direction. Retro-reflection behaviors are obtained when $C_{xy} > 0$.

Ashikhmin-Shirley Model

Ashikhmin and Shirley [2000] presented another energy conserving and reciprocal model, which also allowed to represent anisotropic materials.

The isotropic version of the model that will be used in this thesis is the following one:

$$f_r(\omega_i, \omega_o) = \rho_s \frac{m+1}{8\pi} \frac{(n \cdot h)^m}{(\omega_o \cdot h) max((n \cdot \omega_i), (n \cdot \omega_o))} Fresnel(F_0, \omega_o, h) \tag{2.15}$$

where m models the shape of the specular lobe.

Smooth Surface Model

Two new BRDF models were introduced in Löw et al. [2012]: the Smooth Surface BRDF Model and the Microfacet BRDF model. The gaussian distribution commonly used in BRDF models was found to not describe well measured materials and for that reason the Smooth Surface model and the Microfacet model have a sharper specular peak and a broader tail, enabling a more accurate representation of measured data.

The Smooth Surface BRDF model was derived from optical engineering and is defined by the following equation:

$$f_r(\omega_i, \omega_o) = G'Q'S(||D_P||) \tag{2.16}$$

where $G' = 1$, $Q' = F(\theta, \eta)$, and F is the Fresnel reflectance function proposed by Cook and Torrance [Cook and Torrance, 1982] with the extinction coefficient set to zero:

$$c = cos(\theta) \tag{2.17}$$

$$g = \sqrt{\eta^2 + c^2 - 1} \tag{2.18}$$

$$F(\theta, \eta) = \frac{1}{2}\frac{(g-c)^2}{(g+c)^2}\left\{1 + \frac{[c(g+c)-1]^2}{[c(g-c)+1]^2}\right\} \tag{2.19}$$

where θ is set to $\theta_d = arcsin(\frac{||\omega_{iP} - \omega_{oP}||}{2})$ (ω_{iP} and ω_{oP} being the surface tangent plane components of the incident and outgoing directions, respectively), and η is the Frensel reflectance parameter. $S(f)$ is the following equation:

$$S(f) = \frac{\rho_s}{(1 + Bf^2)^C} \tag{2.20}$$

where ρ_s is the specular reflectance, and B and C are model parameters. D_P is computed as follows:

$$H = \omega_i + \omega_o \tag{2.21}$$

$$D_P = H - (H \cdot n)n \tag{2.22}$$

where H is the unnormalized halfway vector and n is the surface normal.

Microfacet Model

The Microfacet BRDF model is defined by the following equation:

$$f_r(\omega_i, \omega_o) = \frac{S(\sqrt{1-(h\cdot n)})F(\theta_{d,}, \eta)G}{(\omega_i \cdot n), (\omega_o \cdot n)\pi} \tag{2.23}$$

where S is Equation 2.20, F is Equation 2.19, G is Equation 2.10, h is the normalized halfway vector, and n is the surface normal.

2.2.3 Approximation and Error Metrics

Measured materials are commonly approximated with analytical BRDF models in order to improve storage efficiency and to allow importance sampling. Non-linear optimization is used in order to find the parameters of an analytical BRDF model that better reassemble the measured data.

In order to understand which analytic BRDF models best approximate measured BRDF data, the 100 materials of the MERL database were approximated with 7 different analytical BRDF models in Ngan et al. [2005]. This study provided insights into the expressivity of the different analytical BRDF models. A key aspect in the approximation step is the error metric selection. In this case, the objective function used in the optimization step was the minimization of the RMS error metric weighted by the cosine of the incident light direction and the solid angle, in order to compensate for the reflectance increase towards grazing angles and the measurement sampling.

The authors emphasize that the best fit according to their metric does not always correspond to the best visual match, which they found to be highly dependent on scene geometry and illumination.

There is no consensus in the literature about metric selection, and every researcher tends

to apply corrections from their previous experience. In Ngan et al. [2005], the log and cube root compressive metrics were not used as the authors found that the specular highlights became too blurry, so they used the cosine weighted RMS with a solid angle correction term. Similarly, in Lafortune et al. [1997] the cosine weighted RMS metric was used. On the other hand, Matusik [2003] emphasized the need of compression to obtain a good approximation of glossy materials and used a logarithmic function as error metric. A logarithmic metric also provided better approximation in comparison to a cosine weighed RMS metric in Löw et al. [2012].

One of the goals of this thesis is the evaluation of the different error metrics used in the literature, and to try to improve the understanding of which metric leads to a perceptually closer approximation of a measured material.

A method to navigate through a uniform material appearance space was created in Ngan et al. [2006]. The pixel-by-pixel differences between synthetic images generated with different BRDF models were used to create this space. A precomputation step was used to generate all the images used in an interactive interface to aid the material design. This technique would probably give a good performance if used as error metric during the optimization process, but it would require the generation of a synthetic image in each iteration step of the optimization process, making it computationally expensive.

A different approach to approximate real world materials in computer graphics was proposed in Westlund and Meyer [2001], where a virtual goniometer able to simulate a gloss meter and a haze meter was derived. This technique could be used to calculate the gloss and haze values for each set of parameters of a given analytical BRDF model. Then, a look-up-table could be created to relate a set of measurements with a set of BRDF parameters to use at rendering time. The idea of mapping ASTM standards to BRDF parameters is also used in the projection defined in Chapter 5. However, the projection defined in this thesis

makes use of multiple ASTM features at once in order to have a better discrimination across a broad range of materials.

2.3 Perceptual gloss spaces

The description of surface gloss would be facilitated by having a perceptual gloss space in the same manner as the description of color is greatly simplified using CIELAB. Perceptual spaces allow to perform key tasks like easy navigation of the space, or to compute the difference between two materials. Two perceptual spaces have been defined for gloss and are next described.

In Pellacini et al. [2000], a perceptual space of glossy materials represented by the Ward BRDF model was created. Two different experiments were carried out to first find the dimensionality of the gloss space for a set of samples, and then to scale the perceptually meaningful axes found in the space.

In the first experiment, a set of 27 materials was created by varying each of the Ward BRDF model parameters (ρ_d, ρ_s, and α). Then a perceptual experiment was performed in which the observers had to determine the difference between each pair of materials' images. That was used to fill out a proximity matrix, which was input into a multidimensional scaling algorithm (MDS). Two dimensions were determined to be enough to describe the data set.

The next step was to determine a set of perceptual axes that described the two dimensional space found in the first experiment. Perceptual features reassembling contrast gloss (c), which describes the relation between the diffuse component and the specular peak, and distinctness of image gloss (d),which defines the sharpness of the reflections, were found to represent the two dimensions of the space obtained with the MDS process.

The second experiment was used to scale each axis separately. A magnitude estimation experiment was performed, in which the observers had to determine the glossiness of a

material shown on an image. Those images were created by varying ρ_d and ρ_s while fixing α in order to scale the contrast gloss (c), and varying α while fixing the other parameters to scale distinctness of image gloss (d).

Finally, in order to obtain a perceptually uniform gloss space, an analytical equation was derived to relate the observers' responses with a perceptual feature composed of the BRDF parameters. The perceptual parameters were mapped to the Ward BRDF parameters using the following equations:

$$d = 1 - \alpha \tag{2.24}$$

$$c = \sqrt[3]{\rho_s + \rho_d/2} - \sqrt[3]{\rho_d/2} \tag{2.25}$$

$$L^* = f(\rho_d) \tag{2.26}$$

where ρ_d and ρ_s are the diffuse and specular reflectance, respectively, α is the Ward BRDF model parameter, and f corresponds to the CIELAB lightness function, which is related to ρ_d for compatibility.

The results of the experiment were also used to define a perceptual distance metric in the space:

$$D_{ij} \propto \sqrt{[c_i - c_j]^2 + [1.78 \times (d_i - d_j)]^2} \tag{2.27}$$

Finally, the authors rewrote the equations of the space to obtain the Ward BRDF parameters needed given a specific position in the perceptual space:

$$\rho_d = f^{-1}(L) \tag{2.28}$$

$$\rho_s = \left(c + \sqrt[3]{f^{-1}(L)/2}\right)^3 - f^{-1}(L)/2 \tag{2.29}$$

$$\alpha = 1 - d \tag{2.30}$$

where f is the CIELAB lightness function.

This space has multiple applications, for example it can be used to describe differences in apparent gloss, either by using the differences in each dimension or by using the difference equation described in Equation 2.27, another use is to aid the material design by adjusting the perceptually uniform parameters L, c, and d instead of the non-linear parameters of the BRDF model. As a limitation, the authors mention the fact that the space is only valid for the materials used in the study, but they note that it should generalize well.

The main limitation of this work that this thesis will overcome is the fact that the space defined in Pellacini et al. [2000] is only defined for the Ward BRDF model, thus not being able to input measured data into the space. By being able to input both, measured data and analytical BRDF models, at the same time into a space the difference between the materials could be used to guide the optimization process used to approximate a measured material with an analytical BRDF model or to perform quality control applications.

A second perceptual space for gloss was introduced in Wills et al. [2009]. In this case, 55 measured materials from the MERL database [Matusik et al., 2003] were used to define the space. A pairwise comparison experiment was performed in which three images were presented for each trial, and the observer had to select the image that was closer to the one in the center. The results of the experiment were analyzed with the new technique developed by the same authors in which they used MDS to analyze pairwise comparisons [Agarwal et al., 2007], allowing to first estimate the dimensionality of the data, and then to construct the embedding space.

The result from this experiment was a two dimensional space that embedded the 55 materials used in the experiment. Then, the authors studied the correlation of the space obtained with the gloss dimensions defined in Hunter and Harold [1987], the ASTM gloss dimensions mentioned as significant in previous work [Pellacini et al., 2000; Westlund and

Meyer, 2001], the parameters of the 7 analytical BRDF models used to approximate the MERL database in Ngan et al. [2005], and the 2-dimensions described in Pellacini et al. [2000]. A varying degree of success was obtained when correlating each of those dimensions to a direction in the two dimensional space defined in this work. However, the authors didn't define any specific perceptual axes in the space, not allowing to find where a new material would be located, or to compute the difference in gloss perception terms between two materials.

To be able to populate the two dimensional space with materials located in any position, an interpolation technique was also presented in this work. As a preprocess, a delaunay triangulation of the space was calculated, where the location of each material used to create the space was used as a vertex. The goal of a delaunay triangulation is to maximize the minimum angle of every triangle in the triangulation procedure, avoiding skinny triangles that would lead to non smooth transitions in the space. Then, the following process is performed to obtain the material corresponding to a given set of coordinates: First, the triangle that contains the specified coordinates is found. Second, the barycentric coordinates of the point are used to linearly interpolate between the materials located at the vertices. This interpolation process is useful for measured materials, where the measurements itself can be interpolated using the barycentric coordinates as weights. At the same time, if the materials in the vertices were approximated with an analytical BRDF model, the barycentric coordinates could be used as weights to interpolate the parameters of those materials and recreate the two dimensional space. This interpolation strategy is linear for each triangle, but results in an overall non-linear interpolation process.

In summary, neither of the perceptual gloss spaces presented in the literature allows to input both, measured data and analytical BRDF models, into the space. Thus, not allowing to compute the difference between two materials and use this space to perform the

approximation step.

Another limitation of the current work is that the perception of gloss in respect to color has not been studied. The spaces described above only considered grayscale stimuli, taking out the color attribute. The contrast is probably involved in the relation between color and gloss, but more work needs to be done in that area.

Chapter 3

Media dependent gloss perception

Gloss communication, and the higher level material appearance communication is becoming more important every day with the increase in personal manufacturing and the need for the customer to preview a final product while short-runs and cost constrain do not allow the use of hard-copy proofs.

Color appearance models were developed to account for the viewing conditions and its effects on the perception of color. The same color stimuli seen on a hard copy and on a self-luminous display produce different color perceptions. Color appearance models are successfully used to model this changes in appearance and enable to create the same color perception on different media.

It is currently unknown if the same effect is present in gloss perception, or if there is any need for gloss appearance models. This project is designed to study if there exists a fundamental difference in cross-media gloss perception. Gloss communication could be improved with a transformation that accounted for the difference between the representation of a material seen on a display and the real material.

Gloss also varies in other dimensions than color. Vangorp and Dutré [2008] studied the gloss perception dependence on an object's shape, and found that the material appearance

perceived varied depending on the shape of the object. By using the uniform gloss space defined in Pellacini et al. [2000] the authors did model the shape dependence and were able to correct for it, being able to match the gloss appearance of two objects with different shapes.

In this project, in order to understand the gloss perception difference between real objects and synthetic objects seen on a display, three different matching experiments were conducted.

In the first experiment, the observers had to match real objects in a custom-build light booth. This lead to an understanding of the accuracy of the observers and their variability when performing the task with real objects.

In the second experiment, the observers repeated the same task but this time they performed it on a display with synthetic images representing the real objects. As with the previous experiment, this lead to an understanding of the accuracy and variability of the observers performing this task on another media. Furthermore, it enabled the comparison of how accuracy and variability varied from the real objects to the simulations.

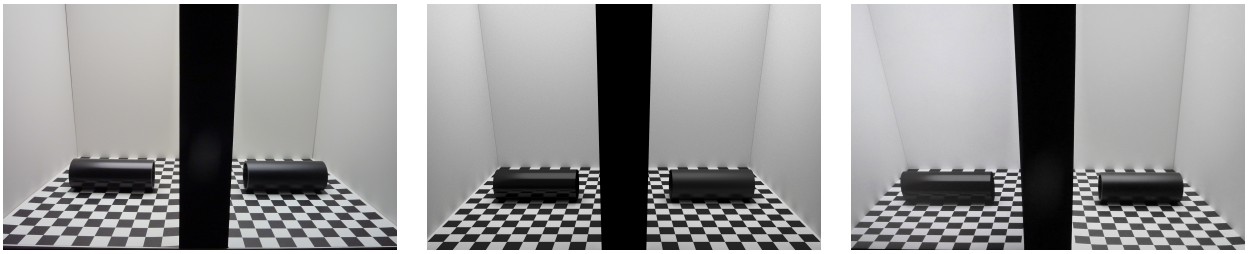
In the third experiment, a cross media matching experiment was performed in which the observers had to select the simulation of an object that matched a real object. This lead to an understanding of the influence of the media used in the matching task.

3.1 Setup

The setup created to perform the experiments can be seen in Figure 3.1. A custom-built light booth is used for the first experiment, and the combination of the light booth and a display are used for the rest of the experiments. The perspective between what is seen on the display (left) and what is seen in the light booth (right) does not match because the synthetic images are generated from a specific viewpoint and have the correct perspective only when the observer is located at a specific position and looking at the samples. Figure 3.2 shows the observer stimuli as seen in the three different experiments with the most and least

glossy sample being presented.

Figure 3.1: Setup used for the experiment. From left to right, 30-inch HP ZR30w display, custom-built light booth, and lazy susan used to provide easy access to the samples for the users.



(a) Real vs. Real (picture) (b) Display vs. Display (image) (c) Display vs. Real (picture)

Figure 3.2: Viewing conditions of the three different experiments, displaying the most glossy sample (left) and least glossy sample (right).

Next, the design decisions and the detailed information for each setup is explained.

Real Scene

To perform the perceptual study on real objects, a scene that is easy reproducible when generating synthetic images and at the same time enhances the material discrimination was designed.

The custom-built light booth and the material samples used for this set of experiments can be seen in Figure 3.1 right. The light booth consists of a wood-structure with an opening of the same size as the 30-inch HP ZR30w display, the one used to display synthetic images, while the depth of the light booth is the same as the vertical edge of the same 30-inch display. A photo studio light source with CFL light bulbs was used to lit the scene. The inner and outer diffuser of the light source were used, and another diffuser was placed on top of the light booth. This light source provided constant chromaticity over angle and the peak luminance of the real scene was slightly lower than the peak luminance of the display ($330cd/m^2$).

The light booth was split vertically in order to accommodate the two scenes used for the matching experiment. Comparisons when the two objects were too close together were avoided by having the regions physically separated, as otherwise it would improve the accuracy of observer judgements to the point that it would be difficult to relate the results from the single media experiments to the cross-media experiment. The separation of the real scenes was equal to the monitor frame plus the light booth frame, in order to allow to have the same distance between samples in all three experiments.

To enhance the material discrimination, a checkerboard pattern was placed on the bottom of the light booth, allowing to see more or less distinctive reflections of the checkerboard pattern depending on the glossiness of the objects used in the experiment.

The object shape selected for this experiments was a cylinder. Several advantages are found on a cylinder over other shapes: they are easy to manufacture, easy to represent

in synthetic images due to their analytical definition, easy to wrap paper around them (see later), and were found to provide a high material discrimination in Vangorp et al. [2007]. Cylinders of diameters 4 and 6 centimeters were created and evaluated. The ones with a diameter of 6 centimeters were finally used as the lower curvature allowed a better discrimination between samples. This was because the spread of the specular lobe and the reflections of the environment were occurring over a larger area and were easier to perceive.

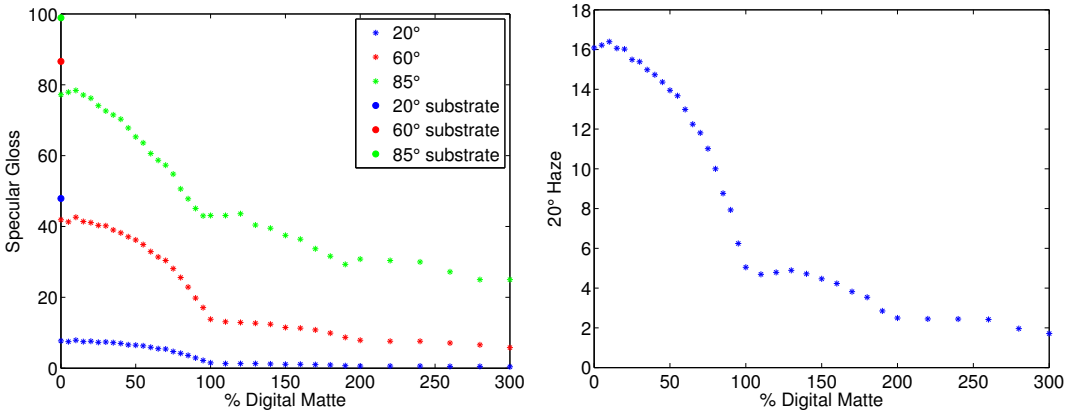


Figure 3.3: From left to right, 20°, 60°, and 85° Specular Gloss, and 20° Haze of the samples used in the experiment, respectively.

The key aspect of the physical setup is the material selection. A set of materials that only varied in gloss was obtained, while keeping other appearance attributes like texture and color constant. In order to achieve this goal, the Digital Matte feature of the HP Indigo 5000 Digital Press was used. The Digital Matte is a varnish that decreases the gloss of the surface on which it is applied. For that reason, the glossy HP Photo Paper was selected as starting point. Then, a first layer of 100% black ink was applied on top of the paper in order to increase the contrast and enhance gloss perception. Then, a varying amount of Digital Matte (0-300%) was applied on top of the black ink in order to obtain 36 samples of different gloss levels. The percentage of Digital Matte for each sample was determined visually to obtain approximately equally spaced samples in terms of perceived gloss. Figure 3.3 shows

the Specular Gloss and the Haze of the created samples measured with a Elcometer 6015 DOI Haze Meter. In spite of also measuring the distinctness-of-image gloss, it is not reported in this document as it it didn't allow to differentiate the samples used, mainly because of its broad specular lobe.

Synthetic images

An accurate representation of the materials, geometrical objects, and lighting of the custom-built light booth needed to be obtained in order to represent the real scene in synthetic images. The scene described before was carefully designed taken into account that it had to be used afterwards for rendering and for that reason simple geometric objects like a rectangular shapes and cylinders were used instead of other selections.

The light source used in the physical scene was carefully measured in order to correctly simulate it. The light source was measured with a PR-650 spectroradiometer at 10° intervals from the normal direction up to 80°. A constant chromaticity over angle was found, and the luminance fall-off measured was approximated with a cubic polynomial. This approximation was then implemented into the rendering engine.

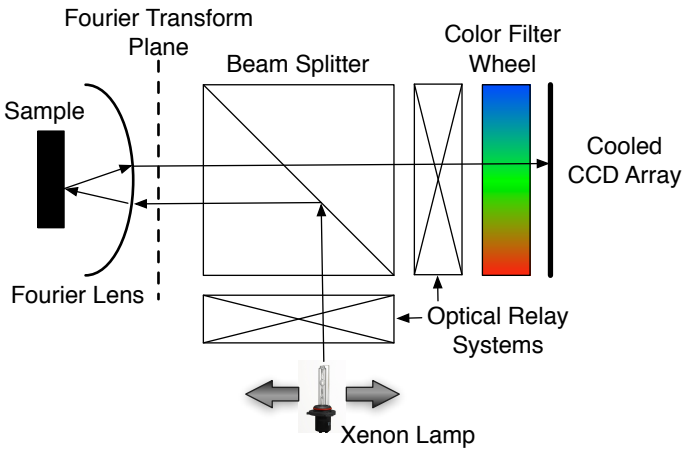


Figure 3.4: Design schematic of the Eldim EZContrast 160R.

Finally, the most important aspect to represent in the synthetic images is the material appearance of the different samples. In order to get the highest angular sampling possible within the available resources, the samples were measured with the Eldim EZContrast 160R at the Centre de Recherche sur la Conservation des Collections in Paris, France. This device allows to measure all the outgoing directions (up to 80° in θ) at once given a single incident direction. Figure 3.4 shows the diagram of the instrument, in which the light reflected from the sample over all the directions is projected onto a plane with a special fourier lens and that plane is imaged by the device sensor. A beamsplitter is used in the middle in order to set the incident light direction, which is projected back to the sample at the selected incident direction using the fourier lens. Five transmission filters are used in front of the CCD sensor to obtain a good approximation of CIE XYZ, and for each measurement a different exposure is performed for each filter.

While all the outgoing directions are measured at the same time, a separate measurement is required for each incident direction. Due to the time required for each measurement and the number of samples to measure, only the following incident directions were measured for each sample: $\theta = 5°, 15°, 30°, 45°, 60°$ and 70° with $\phi = 0°$. It's important to note that the material will be considered isotropic, even thought paper substrates are made of fibers that produce a slight anisotropy. Each of the measurements was later on calibrated using the following equation:

$$f_{sample}(\omega_i, \omega_o) = \frac{i_{sample}(\omega_i, \omega_o) - i_{black}(\omega_i, \omega_o)}{(i_{ref}(\omega_i, \omega_o) - i_{black}(\omega_i, \omega_o))\pi} \tag{3.1}$$

where f_{sample} is the BRDF defined at the direction (ω_i, ω_o). The i_{sampe} is the measurement of the sample, the i_{black} is the measurement with a black trap in front of the measurement port, and the i_{ref} is the measurement of a PTFE created from teflon powder. The i_{black} measurements were required due to the stray light produced by the fourier lens in the mea-

surement device, however it was only measured for incident directions up to 30°, as the stray light is considered to be negligible for higher incident directions.

Two options are present when measured materials are considered for rendering. The first one and more commonly used is to approximate the measurements with an analytical BRDF model and the other technique is to render the measured data directly by means of interpolation. The option to approximate the measurements with an analytical model was selected in this paper. In the future, it would be interesting to explore the direct rendering of the measured data, for example using the technique proposed in Stark et al. [2005].

In order to approximate each material a Lambertian lobe was used to represent the diffuse component, and it was set to the 45:0 measurement, and the parameters of a specular BRDF model were non-linearly optimized. The materials used in this work were well approximated with the Ashikhmin-Shirley BRDF model and the RMS cosine weighted error metric. This selection differs from the results obtained in Chapter 4, in which several specular lobes were needed to faithfully represent the materials of the MERL Database [Matusik et al., 2003]. This difference is probably due to the fact that the Ashikhmin-Shirley BRDF model can better represent the material samples used in this study than the ones of the MERL Database.

The Physically Based Ray Tracer (PBRT) [Pharr and Humphreys, 2010] was used to generate the synthetic images presented to the observers from the specific observers' viewpoints. The rendering is performed after transforming the XYZ tristimulus values of the materials and the light source to the sharpened cone responses obtained with the CAT02 matrix, following the technique described in Ward and Eydelberg-Vileshin [2002].

Finally, the display was characterized using a PR-650 spectroradiometer and the Day method [Day et al., 2004], presenting a good additivity and obtaining a mean $CIEDE2000$ of 0.36 when displaying the colors of the 24 patches of the Macbeth Color Checker.

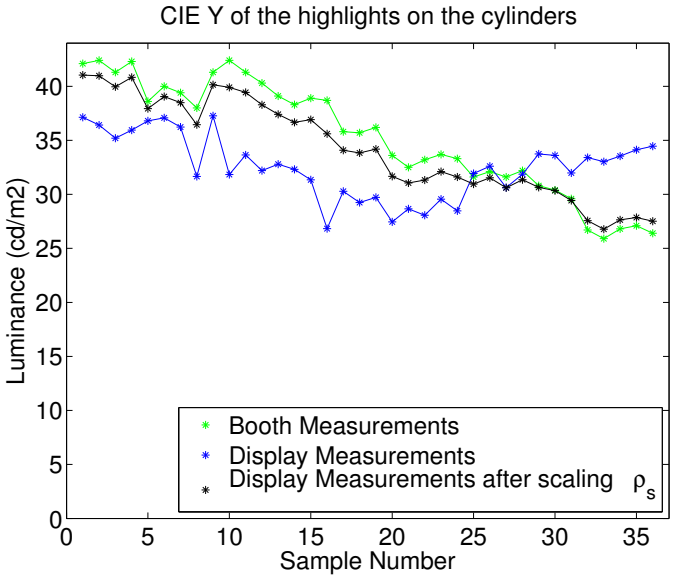


Figure 3.5: Luminance of the highlights on the cylinders measured on the light booth and the display using a PR-650.

In spite of the careful measurements and approximations of the samples and light source, the luminance of the highlights on the cylinders slightly differed when different media were compared (see Figure 3.5). The measurement's accuracy, and specially the approximations of the specular component of the materials are probably the main reasons for observer differences. In order to address this limitation, the specular reflectance (ρ_s) of the materials was scaled to account for the difference between the measurements of the cylinders in the light booth and the cylinders displayed on the monitor. This scaling process improved the representation of the specular peak luminance of the real samples on the simulated images.

3.2 Experiments

The same matching task was performed in the three experiments. The observers were asked to find the match to a sample reference given another full set of 36 samples to choose from.

In each experiment, the observers had to match 8 samples. The samples were carefully

selected to be just noticeably different from the other samples selected to be matched, while keeping some samples on the low and high end.

The three experiments were conducted with the setup seen in Figure 3.1. The experiment was conducted in the dark, in the single media experiments the apparatus not in use was turned off. Fifteen observers with normal color vision and normal or corrected to normal visual acuity participated in the experiments.

Experiment 1: Real vs. Real

In the first experiment, the observers had to match real objects in the custom-built light booth. The reference samples to be matched by the observers were placed, one at a time, in the left side of the light booth. For each of them, the observer had to select the sample that matched the reference sample.

Observers wore latex gloves to avoid damaging the appearance of the samples, as the grease of the skin would rapidly dull the appearance of the paper samples. To browse the 36 material samples easily a lazy susan was built, a circular surface with bearings underneath that allowed to rotate the surface freely (see Figure 3.1 right). A set of dowels were placed along the circle, allowing to set and secure the 36 samples used in the experiment. This setup allowed the observer to efficiently change the sample to be inspected inside the light booth. The samples were sorted from most glossy to least glossy (in terms of % Digital Matte) along the circle.

Before starting the experiment, each observer was trained in how to use the setup. First, the goal of the experiment, to better understand gloss perception in different media was explained. Second, the samples were presented to the observer, telling them that cylinders were wrapped with paper of different gloss levels and that they were sorted from most to least glossy. Then, a demonstration was given by placing the most and least glossy samples

inside the light booth while explaining how to place and align the samples in the small rubber pieces that kept the cylinders in place inside the light booth. Finally, specific instructions to maintain the accuracy of the experiment were given to the observer: only one sample of the matching set was allowed to be taken from the lazy susan at a time, the observer was asked to place the cylinder inside the light booth and rotate it to avoid seeing the seam, to keep the hands off the light booth when making decisions, and to always make the final decision with the sample inside the light booth. Observers were allowed to make a first guess and navigate the gloss range by looking at the reflection seen on the samples in the lazy susan. Once it was clear that the observer understood the setup and the task to conduct, a trial sample that was not recorded was given to the observer in order to let him accommodate to the setup before the start of the experiment.

Results Experiment 1

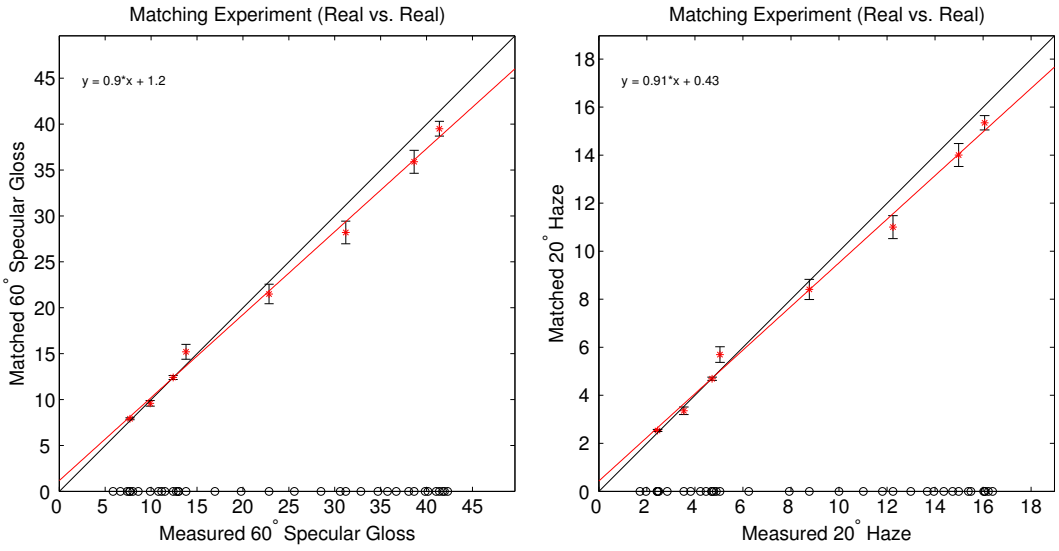


Figure 3.6: Real vs. Real experiment results. Gloss matching performance relative to 60° Specular Gloss and 20° Haze. Sample distribution on the X axis.

The data obtained from the Real vs. Real experiment can be seen in Figure 3.6, in which the mean and standard error matching performance of the observers in respect to 60° Specular Gloss and 20° Haze for each of the 8 samples is shown. The small black circles on the x axis show the measured perceptual properties for all the samples used in the study. The Specular Gloss and Haze measurements of the samples used for this study are highly correlated, thus the matching performance and the fitted linear equation parameters are almost identical. Still, that information is reported in this document for reference. Because of that similarity, the generic term gloss will be used to refer to both perceptual features, Specular Gloss and Haze, when explaining the results.

The black diagonal line shows the 1:1 correspondence if the observers were selecting the sample with equal gloss when performing the matching experiment. It can be seen that for low gloss materials the observers are accurate with their selections and for higher gloss materials the observers tend to select a sample with a lower measured gloss. The fact that the samples are not equally spaced in terms of gloss produces the difference in terms of standard error that is seen across the range. At the same time, the cluster of samples in the high end of the gloss range and the large standard error obtained on that area probably means that the discrimination there was harder than in the low gloss range, or the possibility that the samples in that area are closer perceptually.

A linear equation was fit to the mean observer responses and it can be seen that it correctly models the data, modeling the accurate matching performance for low gloss samples and the gloss underestimation for high gloss samples.

Experiment 2: Display vs. Display

In the second experiment, the observers repeated the same task performed in the first experiment, but in this case the experiment was performed on the monitor by displaying synthetic

images representing the real objects. The synthetic images were rendered from the same point of view from where the observer looked at the real scene in the first experiment, and the camera was tilted down to match the height where the cylinders are located.

In this case, the observers were able to navigate across the gloss range to perform the matching task by using the left and right arrow keys from the keyboard. The spacebar was pressed by the observer when the match was found, and also directed the observer to a black screen indicating the number of samples left to match. As in the first experiment, the sample materials that the observer was able to inspect were sorted by gloss level, from the most glossy to the least glossy, and the system alerted the observer with a visual indication on the screen if they reached the end of the range in a given direction.

Results Experiment 2

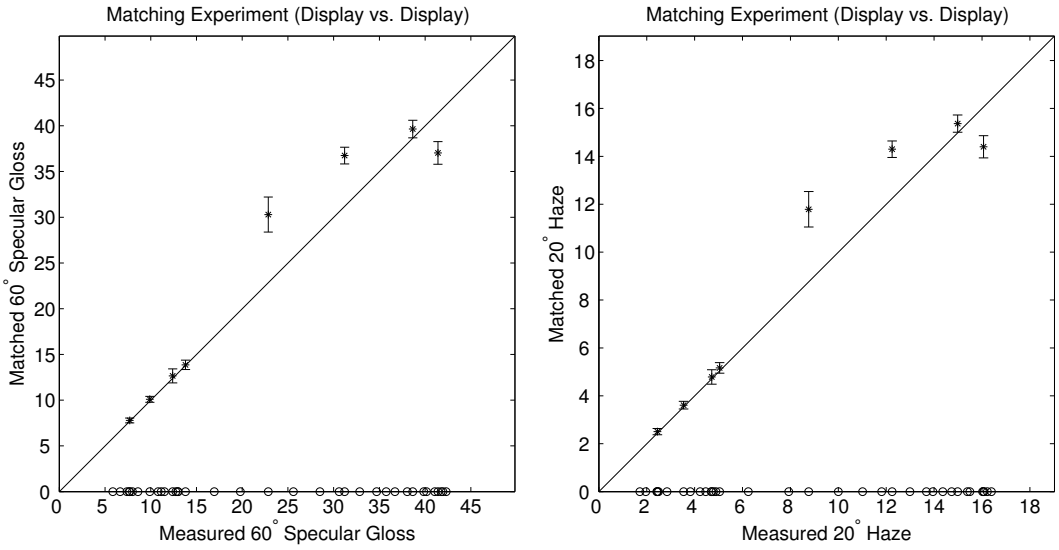


Figure 3.7: Display vs. Display experiment results. Gloss matching performance relative to 60° Specular Gloss and 20° Haze. Sample distribution on the X axis.

Figure 3.7 shows the results obtained from the Display vs. Display experiment. It can be

seen that for low gloss materials the observers are accurate with their selections, samples in the mid range of gloss evaluated are matched by glossier materials, and the glossiest sample was matched to a less glossy material. For the low and high gloss materials of the studied range the observers had the same perception seen as in the Real vs. Real experiment, but the opposite effect was seen on the samples in the middle range.

Experiment 3: Display vs. Real

In the third experiment, the observers repeated the same task performed in previous experiments, but in this case the reference sample was placed in the light booth and the sample the observer had control over was seen on the monitor. In this case, just half of the light booth was visible to the observer while the other part of the scene was physically blocked, at the same time, only a single scene was being shown in the rendered image on the display.

For this experiment, the observer was located between the display and the light booth, and the synthetic images were rendered from that same point of view, and the camera was tilted down to match the height where the cylinders are located.

In this case, apart from different media being evaluated at the same time, there was another major difference. The observers were asked to only use their dominant eye, while closing the other. By doing the experiment with monocular vision, the perspective of the real scene and the synthetic image matched. This might have influenced the experiment, in the same way that binocular cues were eliminated from the second experiment, where no stereo was used. Conducting this cross-media experiment with binocular information in all circumstances would be a challenging experimental design problem, as the technique used to split the image that goes to each eye using glasses would probably also affect the perception of the real scene. For example, the use of polarized glasses would influence and modify the specular reflections seen on the real objects, while shutter glasses would dramatically

reduce the luminance of the real scene to the point that it might become hard to perform the experiment and it could also produce a flickering effect on the real scene. Still, the study of the influence of stereo vision in cross-media gloss perception would be an interesting topic for further research. To evaluate the real life performance of observers the first and second experiments were not run with monocular vision.

Results Experiment 3

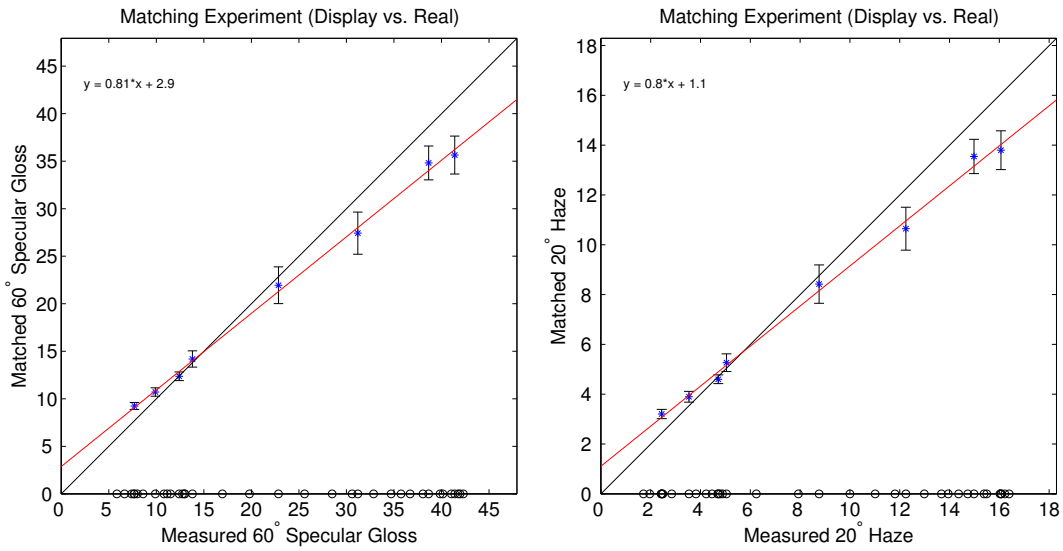


Figure 3.8: Display vs. Real experiment results. Gloss matching performance relative to 60° Specular Gloss and 20° Haze. Sample distribution on the X axis.

Figure 3.8 shows the results obtained from the Display vs. Real experiment. An accurate observers' matching ability is seen for low gloss materials, while a slight gloss underestimation is seen for mid and high gloss materials.

A linear equation was fit to the mean observer responses and it can be seen that it correctly models the data, modeling the accurate matching performance for low gloss samples, with the slight gloss overestimation observed, and the gloss underestimation for high gloss

samples.

3.3 Perceptual Scale

Thurstone's law of comparative judgment (case V) was used to derive interval scales given the data from the psychophysical experiments. The confidence intervals were computed using the empirical formula derived from Monte Carlo simulations of paired comparison experiments in Montag [2006].

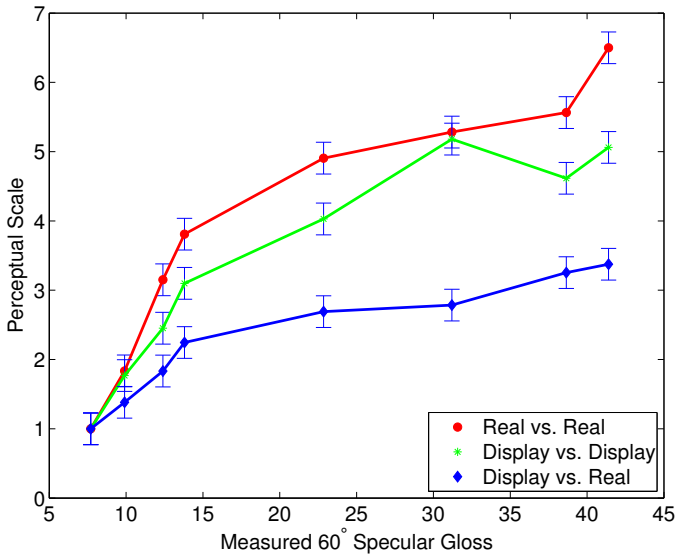


Figure 3.9: Perceptual Scales obtained from the three different experiments.

In pre-testing, samples were chosen by subdivision to form an approximately perceptually uniform scale with sub-JND intervals.

A gaussian distribution was approximated to the observer matching responses for each experiment, and those were later used to compute the probability matrix. For each experiment, every sample was compared against each other and the integration to the cross-over point of the distributions indicated the probability of a given sample to be selected as more glossy. While the probability of the other being selected was 1 minus that probability.

Then, the common Thurstone Case V evaluation was computed using the probability matrix obtained. The χ^2 test was performed and showed that the variance of the samples in each experiment was equal, thus being able to use the Case V.

Figure 3.9 shows the perceptual scales obtained for the different experiments. Significant differences in sensitivity between the different conditions are observed. The highest sensitivity was obtained in the Real vs. Real experiment, a slightly lower sensitivity was observed for the Display vs. Display experiment, and the lowest sensitivity was observed for the cross-media experiment.

Binocular vision, which provides binocular disparity, and the ability of having multiple viewing directions of the samples are probably the main reasons why the highest sensitivity was obtained in the Real vs. Real experiment. Probably, the lack of multiple viewing directions and lack of binocular disparity caused the reduction in sensitivity seen in the Display vs. Display experiment. Finally, the task to perform the experiment using different media decreased even more the sensitivity of the observers in the Display vs. Real experiment, which also had the viewing direction restricted and monocular vision was used.

3.4 Conclusions

In this project, three different gloss discrimination experiments were performed. In the first one, the gloss matching performance with real samples was evaluated. The second experiment evaluated the matching ability when using synthetic images as representations of real objects. Finally, the last experiment evaluated the discrimination ability on the cross-media situation, where real objects were matched with synthetic representations.

The same trend seen in the Real vs. Real experiment was observed in the Display vs. Real experiment, where a high matching accuracy was obtained for low gloss samples and the gloss of mid and high gloss samples was underestimated. A more pronounced gloss

underestimation for those samples was obtained in the cross-media study, as seen with the different slopes obtained.

Similar accuracy was obtained for low gloss samples in the Display vs. Display experiment, but the observer responses for the mid and high gloss samples were the opposite than the ones seen in the other experiments, as gloss was overestimated.

The results obtained in the three different experiments, where the observers have a higher discrimination for low gloss materials and a lower discrimination for medium gloss materials, corresponds with the results found in Obein et al. [2004]. Where the perceptual scaling of gloss using 10 black plates was studied, and it was found that people had more sensitivity to small changes in low gloss samples and people were less sensitive for medium and high gloss samples.

The observers' sensitivity decreased as more constrains were added to the experimental design. The highest sensitivity was observed in the first experiment, where real samples were used. The sensitivity decreased in the matching experiment done with synthetic images due the lack of binocular disparity and fixed viewing direction. Finally, the cross-media study added more constrains like monocular vision and the challenge to deal with multiple media, obtaining the lowest observer sensitivity.

The results obtained in this experiment show that a small increase in gloss might be needed for mid and high gloss materials when synthetic images are used as representations of physical objects. This gloss increase refers to both, Specular Gloss and Haze. The increase of Specular Gloss could probably be implemented by scaling the specular lobe (ρ_s), but there is no direct mapping for performing the same task with Haze. If using the remapped Ward BRDF model presented in Pellacini et al. [2000], the distictness-of-image parameter might be influencing the shape of the specular lobe in a similar fashion as Haze is doing in the samples of this study.

In Section 6.4, the projection of arbitrary BRDFs defined in Chapter 5 will be used to represent the materials of this study in the gloss space defined in Pellacini et al. [2000], and the improved difference equation of the space presented in Chapter 6 will be used to evaluate the trends observed in this experiment.

The gloss range evaluated in this study was limited by the sample creation process used, which limited the highest gloss sample to be in the mid gloss region. As a future work, it would be interesting to evaluate the observer discrimination of high gloss materials, as the observer discrimination is probably going to be different than the one seen on the studied sample set.

Chapter 4

Perceptual evaluation of error metrics

Measured data was rendered directly in Matusik [2003], allowing an accurate representation of real materials. However, this was obtained at the expense of using a complex setup, long measurement times, large amounts of data, and without importance sampling for rendering. Analytical models are commonly used and are more desirable for rendering, as they have a compact representation and allow for importance sampling. For that reason, measured data tends to be approximated with analytical models.

An error metric is used to guide the optimization procedure to achieve the best approximation of a measured material. However, the obtained results do not always correspond with the best visual match because the error metrics currently used do not have any perceptual basis.

The goal of this project is to evaluate which error metric leads to better representations of the measured materials. Two paired comparison studies with reference were performed to evaluate the appropriateness of the different metrics. The reference was a rendered image using measured data and the observers were asked to select the closest approximation to the reference of the two test stimuli presented in each trial. The first experiment compared the combinations of using different BRDF models and error metrics. The second experiment

was performed to understand how close the approximations were to the measured data by adding the measured data to the test set, and by using the error metric that gave the best result in the first experiment.

Multiple factors influence the visual fidelity of the rendering beyond the error metric. The materials approximated, the analytical BRDF models used, the lighting and geometry used in the evaluation scene, and the optimization procedure are key elements involved in the approximation and its evaluation. For that reason, a set of materials and analytical BRDF models were also studied for each error metric and a scene that maximized the material discrimination was created based on previous perceptual studies.

4.1 Stimuli

Ten isotropic materials from the MERL Database [Matusik et al., 2003] were used in this study (*gold metallic paint2, aluminium, blue acrylic, alum bonze, nylon, nickel, blue metallic paint, pearl paint, light red paint, and silver metallic paint*). Those are the same materials as used in Vangorp et al. [2007]. They span different types of materials, colors, and gloss levels. The *copper* material was not found in the database and it was substituted with the *nickel*.

To be able to generalize the fidelity obtained with different error metrics, three analytical BRDF models commonly used in the literature were selected for this study. The Ashikhmin-Shirley [Ashikhmin and Shirley, 2000] and the Cook-Torrance [Cook and Torrance, 1982] BRDF models were selected because they are widely used and also provided the best performance in Ngan et al. [2005]. The Ward BRDF model [Ward, 1992] was also used in this study, due to its wide use in vision science and perceptually based material modeling experiments [Pellacini et al., 2000; Fleming et al., 2003; Vangorp et al., 2007].

Three error metrics used in the literature were evaluated in this study: the root mean

square error (RMS), the RMS weighted by the cosine of the incident direction, and the cube root of the cosine weighted metric:

- Root mean square (RMS)

$$E = \sqrt{\frac{\sum \left(M(\omega_i, \omega_o) - A(\omega_i, \omega_o, p)\right)^2}{n}} \tag{4.1}$$

- Cosine weighted RMS

$$E = \sqrt{\frac{\sum \left(M(\omega_i, \omega_o) cos\theta_i - A(\omega_i, \omega_o, p) cos\theta_i\right)^2}{n}} \tag{4.2}$$

- Cube root cosine weighted RMS

$$E = \sqrt{\frac{\sum \left(\left(M(\omega_i, \omega_o) cos\theta_i - A(\omega_i, \omega_o, p) cos\theta_i\right)^2\right)^{1/3}}{n}} \tag{4.3}$$

where the difference between the measured BRDF M and the approximation A obtained using a given BRDF model with the parameters p is computed across the n pairs of incident and outgoing directions.

The RMS is the simplest error metric, in which the distance between each of the points of the measured data and the approximation obtained with the analytical BRDF model is computed. The weighting factor used in the cosine weighted RMS is added to compensate for the reflectance increase towards the grazing angles when the incident direction goes from the normal direction at 0° to 90° in θ_i. RMS metrics tend to overemphasize the importance of the BRDF peaks in the mirror direction and de-emphasize the off-peak values. For that reason, the empirically derived cube root metric is sometimes used to try to correct for this effect.

There is no consensus in the literature about metric selection, and every researcher tends to apply corrections from their previous experience. In Ngan et al. [2005], the log and cube root compressive metrics were not used as the authors found that the specular highlights became too blurry, so they used the cosine weighted RMS with a solid angle correction term. On the other hand, Matusik [2003] emphasized the need of compression to obtain a good approximation of glossy materials and used a log function as error metric. The cube root metric was used in this paper as the log function behaves badly near zero, and as it has also been used in the literature.

Finally, a scene that maximized material discrimination was defined. The Eucalyptus Grove light probe from Paul Debevec was used in this study because it was found to be the environment map with real world statistics providing the best material discrimination in Fleming et al. [2003]. This light probe also allows to evaluate the color of an object without the need to perform chromatic adaptation. The influence of shape on the perception of material reflectance was studied in Vangorp et al. [2007], where the ability to discriminate if two different geometric objects had the same reflectance or not was analyzed in a psychophysical experiment. The fact that every 3D modeling application uses a sphere as a sample material was one of the reasons driving this work, and the authors found that the sphere was one of the least discriminating shapes for judging materials. One of the shapes that gave the best discriminating accuracy was a blob-like shape, which contained both concave and convex regions. This blob-like shape was selected for this study in order to maximize the material discrimination.

A core task of this project was the fitting process, in which the parameters of a BRDF model are optimized to minimize a given error metric for a given material. The use of a single specular lobe did not provide a good approximation for most of the materials evaluated, and for this reason a two lobe representation was used. The analytical BRDF model parameters

are highly non-linear and the result obtained depends on the initial values used for the optimization process. The analytical form used to approximate measured materials was the following:

$$K = \rho_d \; diffuse + \sum_{i=1}^{2} \rho_s \; specular(p) \tag{4.4}$$

where ρ_d is the diffuse reflectance (RGB scalars), $diffuse$ is a Lambertian lobe, *specular* is a particular analytical BRDF model, ρ_s is the specular reflectance (RGB scalars), and p are the parameters of the *specular* analytical BRDF model. Note that the same analytical BRDF model is used for both specular lobes, but each specular lobe has different ρ_s and p. The non-linear optimization process was first performed with one specular lobe, and then the process was repeated while adding the second specular lobe in order to improve the stability.

4.2 Experiments

The perceptual fidelity of images created using different combinations of materials, models, and metrics were evaluated by performing a two-alternative forced-choice (2AFC) psychophysical experiment with reference.

The reference was a rendered image using measured data, and the observer was asked to select the closest approximation to the reference of the two stimuli presented in each trial. The interface used for the experiment can be seen in Figure 5.3.

The first experiment compared the visual fidelity of each possible combination of the error metrics (RMS, cosine weighted RMS, and cube root cosine weighted RMS) and the analytical models (Ward, Ashikhmin-Shirley, and Cook-Torrance) to the reference.

The reference image was included in the trial selection in a second experiment in order to evaluate the distance between the approximations and the measured data. In this case, the approximations obtained with the three analytical BRDF models using the error metric

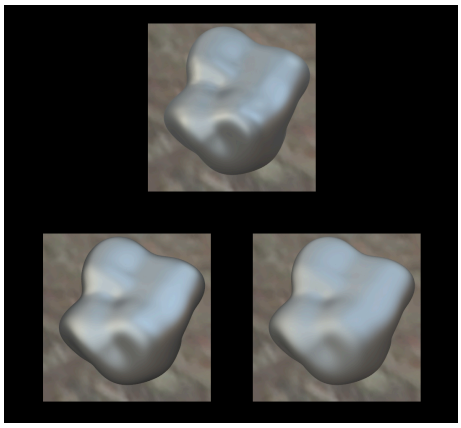


Figure 4.1: Interface used for the 2AFC experiment with reference, developed with Psychtoolbox. The reference image is shown on the top, and the two approximations are shown at the bottom.

that gave the best result in the first experiment and the measured data were compared to the reference image. The camera position on the reference image was rotated 15° around the object to avoid pixel-by-pixel comparisons by the observers.

4.3 Results

The renderings for 7 of the 10 materials can be seen in Figure 4.2. The use of a compressive metric (i.e. cube root) seems to improve the approximation of high gloss materials over the RMS based metrics (see bottom row of Figure 4.2). For low gloss, all the metrics seem to produce similar renderings. The RMS based metrics seem to overfit the specular lobe for high gloss materials. The blue acrylic material was not well approximated with any combination of error metrics and models, the diffuse component was well approximated with the compressive metric, but the specular peak was missed, while the RMS based metrics overfitted the specular lobe.

BRDF plots are commonly used to evaluate the goodness of an approximation, where the in-plane measured and approximated data are displayed. A disconnect exists between BRDF

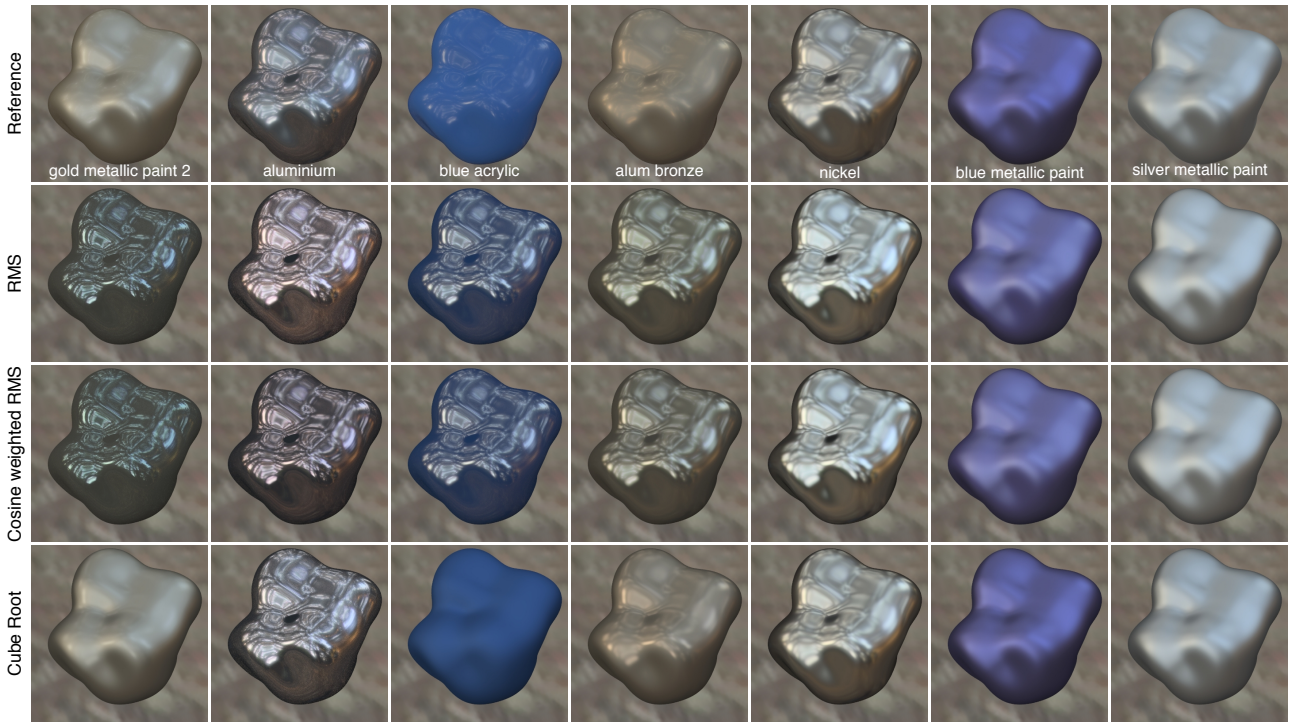


Figure 4.2: Reference and approximations obtained for 7 of the 10 materials used in the study using the Ward BRDF model. A better visual fidelity is obtained with the cube root error metric for high gloss materials, while the RMS based metrics seem to over fit the specular lobe. For low gloss, all the metrics seem to produce a similar visual fidelity.

plots and the visual fidelity of an approximation. Figure 4.3 shows the rendered images and the BRDF plots of a reference material and two approximations. If the BRDF plot is used to evaluate the goodness of the approximation, the cosine weighted RMS metric approximation would be selected as best. However, by looking at the rendered images, it's clear that the metric providing the best visual fidelity is the cube root, in spite of the differences seen in the BRDF plots.

Thurstone's law of comparative judgment (case V) was used to derive interval scales given the data from the psychophysical experiments. The confidence intervals were computed using the empirical formula derived from Monte Carlo simulations of paired comparison experiments as in Montag [2006].

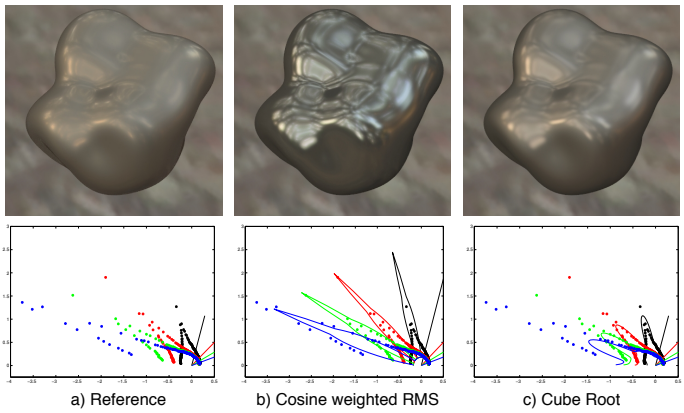


Figure 4.3: From left to right, alum-bronze reference material, cosine weighted RMS, and cube root approximations using the Ward BRDF model. The second row shows the cube root compressed BRDF plots with the measured data and its approximations for the green channel and given different incident directions. The evaluation of an approximation using only the BRDF plots can be misleading.

The interval scales obtained for the first experiment with each material and error metric given a different BRDF model are shown in Figure 4.4. For the Ward BRDF model (Figure 4.4a), the cube root metric is always preferred by the observers, and the sharper the specular lobe (low DOI angle), the more beneficial the use of the cube root metric is. Once the specular lobe broadens, the benefit of using the cube root metric decreases, but still better visual fidelity is perceived by the observers when this metric is used. Without being significant, a small benefit is observed if the cosine weighted RMS metric is used over the RMS metric for the Ward BRDF model. The confusion seen in the blue-acrylic material could be explained with different criteria among observers, where some observers probably gave more weight to the highlights and others to the diffuse component (see Figure 4.2).

The scalings obtained for the Ashikhmin-Shirley and Cook-Torrance BRDF models are shown in Figures 4.4b and 4.4c, respectively. For those BRDF models, the cube root metric is almost always preferred by the observers. However, this is not the case for all the materials and some crossovers appear for the lines connecting the error metrics performance. The main reason of the crossovers is the high number of *outliers* obtained through the optimization

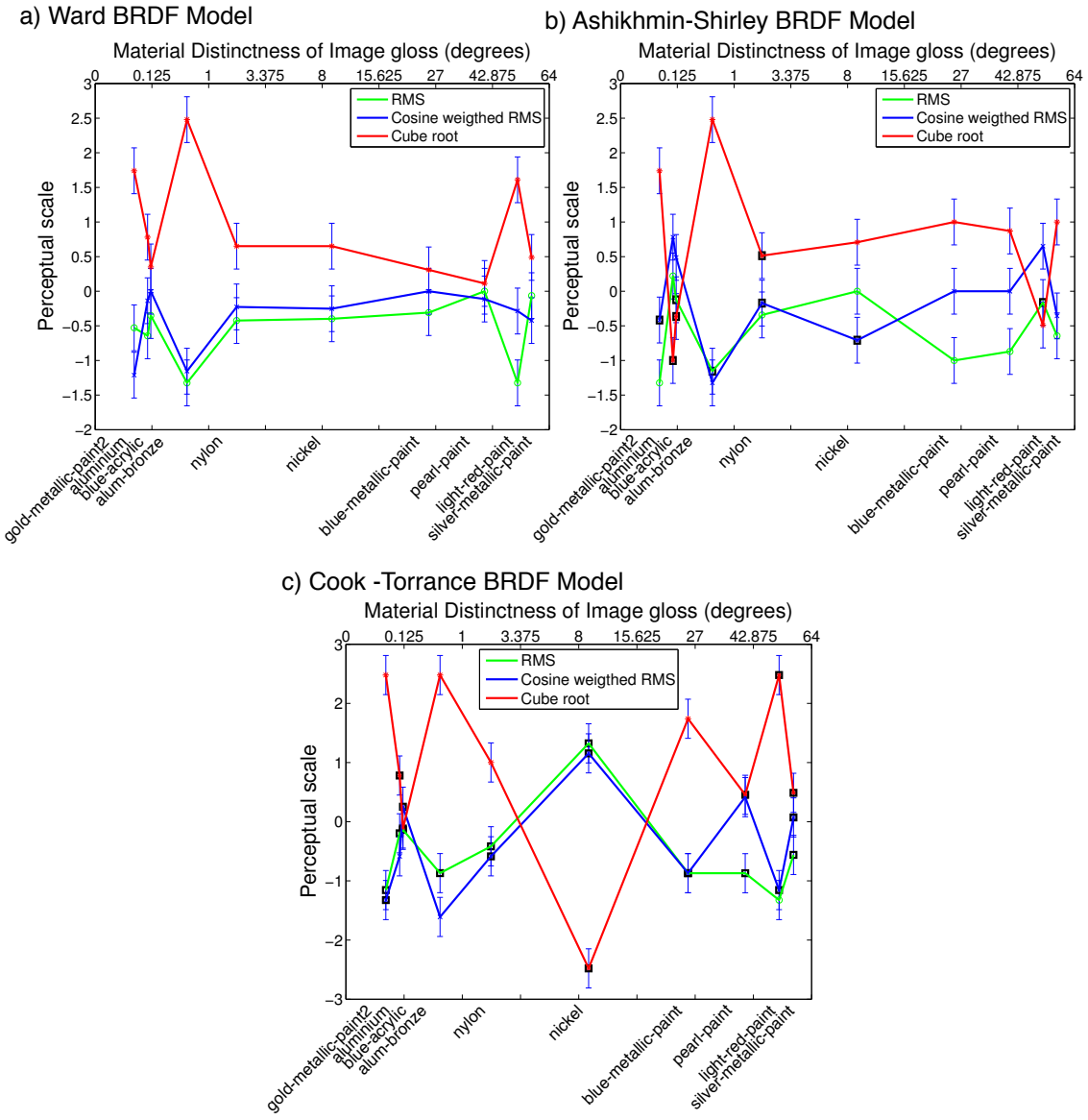


Figure 4.4: Error metric interval scaling across materials sorted by increasing DoI angle for the a) Ward, b) Ashikhmin-Shirley, and c) Cook-Torrance BRDF models.

procedure for those BRDF models. Outliers are those approximations in which a local minimum was reached by the optimization procedure, and are represented with a black square. The reason why outliers were found for the Cook-Torrance and the Ashikhmin-Shirley BRDF models is probably because those BRDF models have two parameters that

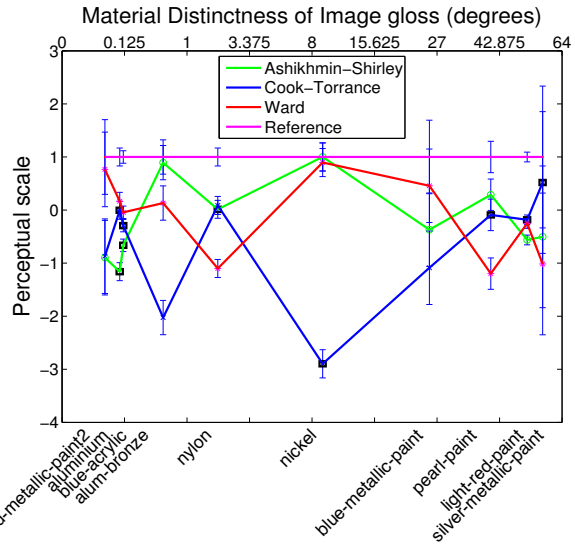


Figure 4.5: BRDF models and measured data (Reference) interval scaling across materials sorted by increasing DoI angle for the cube root error metric.

need to be optimized for each specular lobe making the optimization process less stable, while the Ward BRDF model has only one parameter to optimize.

The cube root metric was selected for the second experiment as it was found to give the overall best approximations in the first experiment. The measured data and the three BRDF models using that metric were compared against the reference image. The interval scales obtained are shown in Figure 4.5. The approximations obtained using the cube root error metric were confused with the measured data for 5 of the 10 materials studied.

4.4 Conclusions

The key finding of this project is the higher visual fidelity obtained using the cube root metric compared to the RMS based metrics for the studied materials. The improvement in visual fidelity using the cube root metric compared to the RMS based metrics is higher for sharp specular lobes and decreases as the specular lobe broadens.

The better performance of a compressive metric can be related to perception, where a similar compression is applied to the lightness channel in CIELAB, and tone mapping operators compress HDR images to be displayed on low dynamic range displays. It would be interesting to repeat the experiment using a high dynamic range display, as it is known that limiting the image dynamic range does change the apparent gloss of surfaces depicted in images [Phillips et al., 2009].

In spite of using two specular lobes to approximate each measured material, only 50% of the materials evaluated were confused with the renderings of the measured data. This fact acknowledges the challenge to faithfully reproduce all the nuances of real world materials using analytical models. It is important to note that the scene used in this experiment was designed to maximize the material discrimination, meaning that in more forgiving scenes the approximations would probably be perceived as being closer to the measured data.

After this project was completed, I became aware of the work described in Löw et al. [2012]. In this paper, the authors first studied the measured data from the MERL database and found that the BRDF models commonly used do not well approximate the shape of the measured BRDF, specially at the tail of the specular lobe. Most BRDF models use a gaussian function, which does not describe the measurements well. The authors presented two new models, one microfacet-based and another one derived from optical engineering. Those models have a sharper peak and broader tail, which allows them to improve the approximation of measured data of the MERL database. Finally, they approximate a set of materials from the MERL database with those two new models using a cosine weighted metric and a log weighted metric. They found that the log weighted metric produced a better perceptual approximation of the measured materials than the cosine weighted metric.

The conclusions found in Löw et al. [2012] in respect to the metric selection correlate with the results obtained in the perceptual study explained in this chapter, stating that the

use of compressive metrics lead to a better perceptual approximation while RMS metrics tend to overfit the specular peak. The use of the new BRDF models in the perceptual experiment performed in this chapter would probably lead to a better representation of measured materials, hence improving the percentage of approximations that are confused with the measured data, while at the same time maintaining the conclusions found in this chapter in respect to he metric selection. The validation of this hypothesis and the evaluation across an even broader set of materials would be an interesting avenue of future work.

Chapter 5

Gloss space projection for arbitrary BRDFs

Realistic material appearance modeling and rendering is an important area in computer graphics, with many applications in areas such as movie industry, advertising, video games, and virtual reality.

A data-driven approach has successfully been used in order to improve the representation of real materials, but at the expense of high angular resolution measurements and storage efficiency. To overcome those limitations, material measurements are commonly approximated with analytical models. Unfortunately, the error metrics used to perform this optimization procedure are not based on perception and do not always lead to the best visual approximation of a measured material.

In this chapter, the use of a perceptually uniform gloss space to perform material approximations is proposed by introducing a projection for arbitrary isotropic BRDFs into the uniform gloss space described in Pellacini et al. [2000]. This projection is based on using ASTM standards, which are used to describe the material appearance of glossy surfaces (e.g

distinctness of image, haze, bloom, etc.), to obtain the *Lcd* coordinates of a material in the gloss space described in Pellacini et al. [2000]. As the perceptual uniform space is defined using the Ward BRDF model, the projection into the gloss space is at the same time an approximation using that model. If other models are also needed, the perceptual distance metric defined in the uniform space can be used to guide the optimization procedure.

The perceptual uniformity of Pellacini's space is also studied. The perceived distance around five positions in the gloss space is evaluated in a psychophysical study to understand the perceptual uniformity of the space. The space was found to be perceptually non-uniform outside the samples used when the space was created, and improvements to the distance metric of the space might be needed to improve the space uniformity.

The current space's distance metric does not allow relating differences in lightness to the other dimensions: contrast gloss and distinctness of image gloss. The lightness perception uniformity of the space was first evaluated in a psychophysical study, where the observers' lightness discrimination was found to decrease as lightness increased. Then, a function was derived to model the lightness perception observed and it was included into the distance metric of the space defined in Pellacini et al. [2000]. It's important to note that the gloss perception using a common LDR display is studied in this work, and similar experiments would need to be performed to study the gloss perception of physical objects or when using HDR displays.

In summary, the main contributions of this project are:

- The inclusion of L^* to the gloss distance equation of the space defined in Pellacini et al. [2000].

- The validation of the gloss space uniformity, finding that the space is non-uniform outside the samples used to create the space in Pellacini et al. [2000].

- The creation and perceptual validation of a projection for arbitrary isotropic BRDFs into the gloss space defined in Pellacini et al. [2000].

5.1 BRDF Projection

In this section the projection of a BRDF into the gloss space defined in Pellacini et al. [2000] is described. This projection is applicable to either measured data, when the required measurements are available, or to analytical models. Then, the perceptual distance between two BRDFs can be computed by using the distance metric defined in the space. As the gloss space is only defined for grayscale materials, materials used in this project will be converted to grayscale.

The projection is based on the relation found between ASTM gloss standards and the uniform gloss space defined in Pellacini et al. [2000]. The ASTM gloss standards of distinctness of image gloss (DOI), bloom, haze, and diffuseness are described as the ratio of light reflected at 0.3°, 2°, 5°, and 15° away from the mirror direction to the light reflected at the mirror direction at 30°. Here, an additional feature at 10° away from the mirror direction was included in order to better sample the specular lobe changes in between haze (5°) and diffuseness (15°).

Figure 5.1 shows the ASTM gloss standard value (color coded from 0 to 1) for each material defined in the *cd* plane at $L^* = 0.5$. The plots for each ASTM gloss standard are obtained by first converting the *Lcd* position of each material in the plane to parameters of the Ward BRDF model by using Equations 2.28-2.30. Second, the Ward BRDF model is evaluated at $\theta_i = 30°$ and $\theta_o = \{30°, 30.3°, 32°, 35°, 40°, 45°\}$. Finally, the ASTM feature for each position in the *cd* plane is computed as the ratio of the Ward BRDF model evaluated at the outgoing angle described by the ASTM feature (e.g. $\theta_o = 30.3°$ for DOI) to the mirror direction ($\theta_o = 30°$).

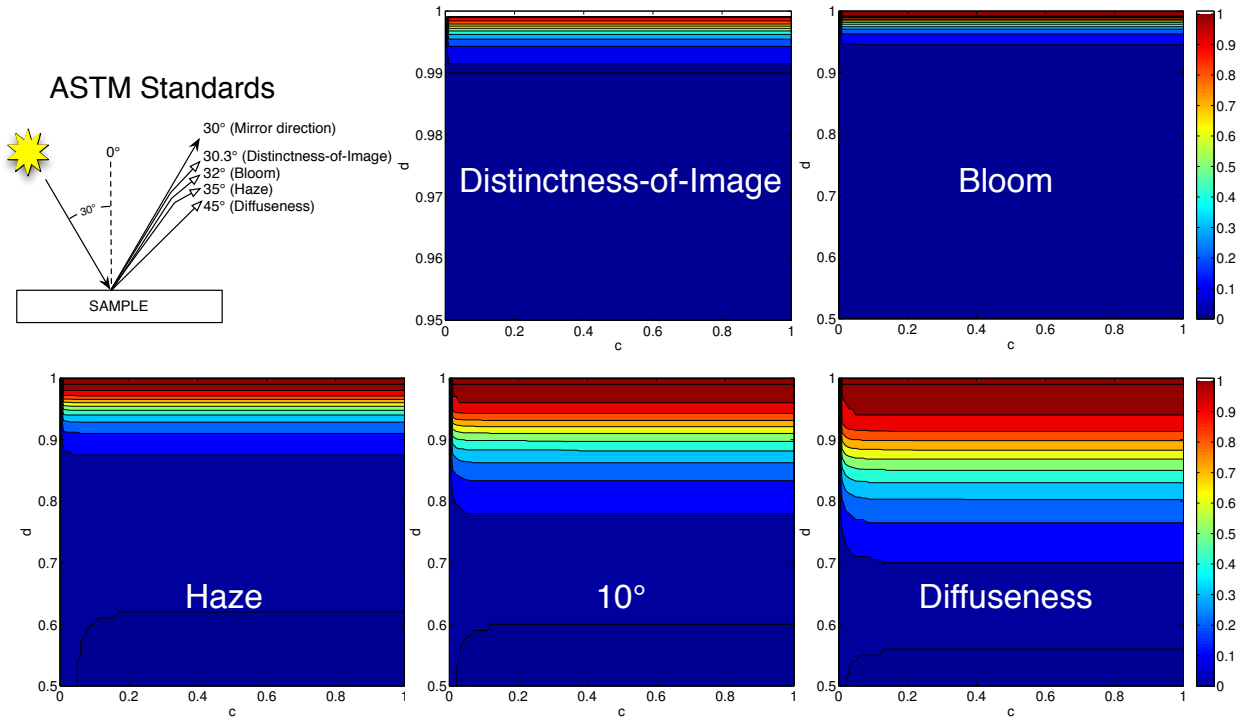


Figure 5.1: Geometry of the DOI, Bloom, Haze, and Diffuseness ASTM standards, and plots showing the ASTM standard value of each point in the cd plane at $L^* = 0.5$. The main idea behind the projection presented is shown in the plots, where each ASTM feature varies in a different region of d while being almost independent of c. Note that the d axis has a different scale for the Distinctness-of-image ASTM feature, as it only varies for materials with a d close to 1.

The key insight in the relation between the ASTM standards and the gloss space is that the different ASTM standards vary in different regions of the d dimension (note different y-axis scale for DOI in Figure 5.1), while at the same time being almost independent of the c dimension (see horizontal lines in Figure 5.1). The ASTM standards sample angles at different distances from the mirror direction and the d dimension models the width of the specular lobe, as sharp specular lobes have a d close to 1 and the lobes broaden as d decreases.

The ASTM standard to d relation also depends on lightness. For $L^* = 0$ complete independence from c exists, as c would only depend on ρ_s (see Equation 2.25), and in that

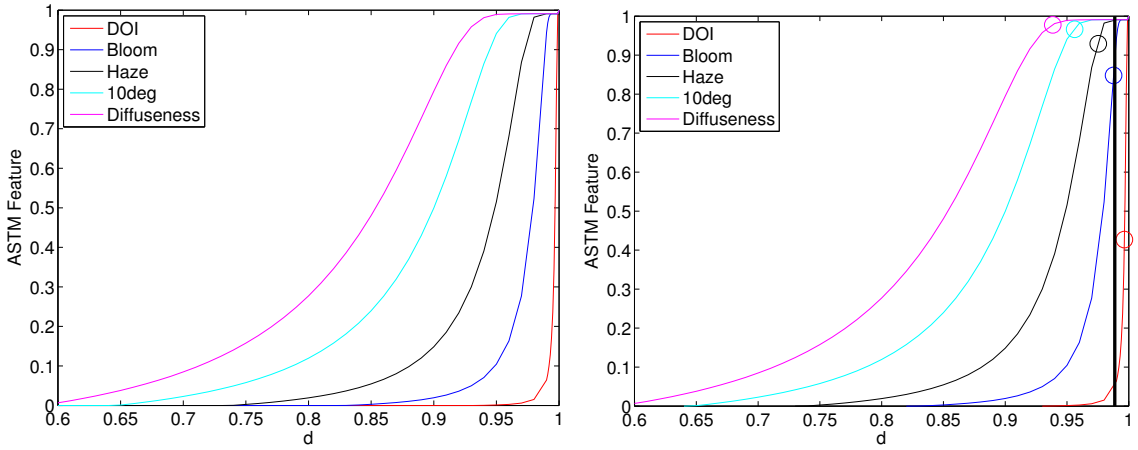


Figure 5.2: Functions relating the ASTM standard features to the Pellacini's d parameter (left). ASTM features (O) and d parameter (black vertical line) obtained for the alum-bronze material of the MERL database (right).

case the lines for each plot in Figure 5.1 would be horizontal. For higher values of L^*, the effect seen for $c < 0.15$ appears to haze, 10°, and diffuseness. In order to have some of that effect present in the projection a $L^* = 0.5$ was selected, higher lightness levels showed a small increase of that effect. For each of the ASTM standards, the mean across each d position was computed across the c dimension to create the curves seen in Figure 5.2, because of this decision the accuracy of the projection to d for low contrast materials will be lower than for other regions.The function is then limited to the [0,1] range.

The relation found between ASTM standards and d is similar in concept to the relation between specular gloss and perceived gloss, where the angle that provides the best discrimination for a given material is used (20, 60, or 85). The same effect is seen in Figure 5.2, where a different ASTM standard provides the best discrimination for a different range of d. As the d of a material of interest might fall in between different ASTM standard functions, d will be obtained by weighting the different ASTM features and giving a higher weight to the ones that provide a higher discrimination.

The first step to obtain the projection of any BRDF into the Lcd space is to compute the

ASTM standard features for the BRDF of interest. Second, the functions of Figure 5.2 are used to obtain d given the weighted ASTM features of the sample. A higher weight to the ASTM features closer to 0.5 is given as it corresponds to the linear region of the function and to the width of the specular lobe at half its magnitude:

$$d_{ASTM_x} = f_x(ASTM_x) \tag{5.1}$$

$$Weight_x = \begin{cases} ASTM_x & \leq 0.5 \\ |ASTM_x - 1| & \text{otherwise} \end{cases} \tag{5.2}$$

$$d = \sum_{x=1}^{5} \frac{d_{ASTM_x} \times Weight_x}{\sum_{y=1}^{5} Weight_y} \tag{5.3}$$

where x is each of the ASTM standard features (DOI, haze, bloom, etc.), $ASTM_x$ is the ASTM standard x for a given sample, and f is the function of Figure 5.2 for each ASTM standard.

Once d is known, ρ_d is obtained directly from the 45:0 measurement geometry (incident:outgoing direction), being commonly used in colorimetry and corresponding to the diffuse reflectance. Next, α is obtained using the following equation:

$$\alpha = 1 - d \tag{5.4}$$

Then, the following equation is used to obtain ρ_s:

$$\rho_s = \frac{M(\omega_i, \omega_o) - \rho_d/\pi}{Ward(\omega_i, \omega_o, \alpha)} \tag{5.5}$$

where M is the BRDF measurement, $Ward$ is the BRDF model of Equation 2.12, and (ω_i, ω_o) corresponds to the 30:30 geometry as it was found that scaling the peaks provided

the overall best perceptual approximation. The 100 materials of the MERL database were approximated using different scaling techniques and scaling the peaks provided the best visual approximation across materials. Finally, Equations 2.24-2.26 are used to obtain c and d given ρ_d, ρ_s, and α.

The presented projection exploits the limitation of the Ward BRDF model of not modeling the Fresnel effect, by just requiring measurements at a single incident direction $\theta_i = 30°$ a part from the diffuse measurement. However, this is at the same time a limitation as the behavior of the material at $\theta_i = 30°$ will determine the appearance at other incident directions. An avenue of future work could be to use several incident directions and weight them to determine ρ_s or the ASTM features, a single incident direction has been used here for simplicity.

Another limitation of the projection is also inherent to the BRDF model used in the *Lcd* space. Commonly used BRDF models, including the Ward model, do not always well approximate measured materials, specially the tails of the specular lobe [Löw et al., 2012]. Two new BRDF models with a distribution more similar to measured data were presented in Löw et al. [2012], and by having a sharper peak and a broader tail they were able to improve the approximation of measured data. The limitation of the Ward BRDF model to accurately represent some measured materials will be an inherent limitation to the projection, as certain materials will not be successfully represented in the *Lcd* space.

5.2 Experiments

Three experiments were performed in order to (1) Determine perceptual lightness differences in the *Lcd* space, to (2) Verify gloss uniformity of the *Lcd* space, and to (3) Validate the gloss projection defined in the previous section.

The method of constant stimuli was used together with a two-alternative forced choice

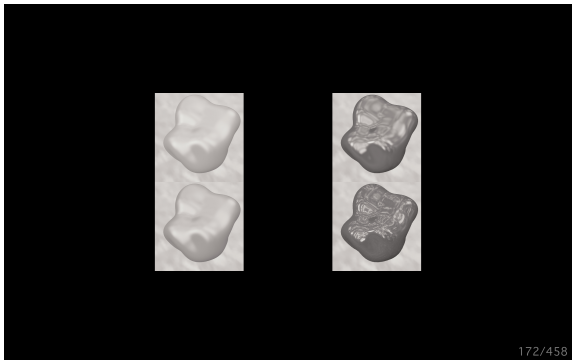


Figure 5.3: User interface of the experiment developed with Psychtoolbox, showing the two pairs of images to be compared.

(2AFC) design. Figure 5.3 shows the interface used in the different experiments, where two pairs of images were presented for each trial to the observers. The standard pair of images was on the left, and the test pair was on the right. The test pair was composed of a given gloss center and a test sample. The upper/lower position of the images for the test pair were randomized for each trial. All the test samples for each gloss center were randomized for each observer, and the gloss center sequence was kept constant.

The question that the observers had to answer for each of the trials was the following one: *Which pair of images is more similar?*. Observers were instructed to judge the overall material appearance difference, i.e. taking into account both lightness and gloss differences. The left and right arrow keys on the keyboard were used to select the pair of images considered to be more similar. Twenty-one observers with normal color vision and normal or corrected to normal visual acuity participated in the experiment. The experiment was conducted in the dark and each observer performed a total of 698 judgements between the three experiments.

In order to maximize the material discrimination in the experiment a blob-like shape and the Eucalyptus Grove light probe from Paul Debevec were used. This geometry gave the best material discrimination accuracy in Vangorp et al. [2007] and this environment map was found to be the environment map with real world statistics providing the best material

discrimination in Fleming et al. [2003], respectively.

The 30-inch HP ZR30w display was used for the experiment and was characterized using a PR-655 spectroradiometer and the Day method [Day et al., 2004]. A good display's additivity was observed, and a mean $CIEDE2000$ of 0.33 was found when displaying the colors of the 24 patches of the Macbeth Color Checker.

The Physically Based Ray Tracer (PBRT) [Pharr and Humphreys, 2010] was used to generate the synthetic images presented to the observers. The pfstools framework [Mantiuk et al., 2007] was used to apply the Reinhard et al. [2002] global tone mapping operator (key= 0.18 and $\phi = 1.0$) to a tiled image containing all the images used for each experiment. The Reinhard et al. [2002] tone mapping operator was selected due its simplicity and its widespread usage in the computer graphics literature, and it is computed with the following equation:

$$\bar{L_w} = exp\left(\frac{1}{N}\sum_{x,y} log(\delta + L_w(x,y)\right) \tag{5.6}$$

where $\bar{L_w}$ is the key of the scene (subjective assessment of a light, normal, or dark scene), $L_w(x,y)$ is the luminance of the pixel (x,y), N is the number of pixels in the scene, and δ is set to 0.00001 to avoid the singularities of the log function. Then, the scaled pixel luminance is computed:

$$L(x,y) = \frac{\alpha}{\bar{L_w}} L_w(x,y) \tag{5.7}$$

where α is set to 0.18. Finally, the last step in the tone mapping operator is to compress the highlights:

$$L_d(x,y) = \frac{L(x,y)}{1 + L(x,y)} \tag{5.8}$$

After applying the tone mapping operator, the resulting LDR image was linearly mapped to the display's dynamic range. Next, the color for each pixel was set to be neutral by scaling its X and Z components: $X = Y\frac{X_{wp}}{Y_{wp}}$, $Z = Y\frac{Z_{wp}}{Y_{wp}}$, being XYZ_{wp} the display's white point.

Finally, the inverse of the display model was used to convert from XYZ to display digital counts.

5.2.1 Experiment 1: Adding lightness to the space's distance metric

The goal of the first experiment is to add a function for L^* to the gloss distance measure seen in Equation 2.27.

The *Lcd* space difference equation presented in Pellacini et al. [2000] does not define how the differences in lightness are related to differences in the other dimensions of the space, c and d. For that reason, differences between materials with different lightness can't be resolved using the current equation.

The experimental design explained in the previous section was used to understand and model the behavior of L^* in the *Lcd* space. The standard pair was defined to only vary in a single dimension ($d = \{0.9, 0.9281\}$), while having constant lightness and contrast ($L^* = 30$, $c = 0.25$). The distance between the standard pair was equal to $D_{i,j} = 0.05$ in the uniform space described in Equation 2.27.

Several gloss centers were studied to better understand the perception of lightness differences across the space. Five positions in the *cd* plane were selected to study if there was a lightness perception difference depending on the position on that plane ($cd = \{0.2, 0.85\}$, $\{0.3, 0.9\}$, $\{0.4, 0.95\}$, $\{0.1, 0.8\}$, $\{0.15, 0.925\}$). Then, a gloss center was defined at $L^* = \{10, 20, 40, 60, 80\}$ for each of the *cd* positions.

Eleven test samples were defined for each gloss center, varying only in L^*. They spanned the range $[L^* - 10, L^* + 10]$ with an interval of 2 when the gloss center was at $L^* = 10$, and the range of $[L^* - 20, L^* + 20]$ with an interval of 4 for gloss centers defined at other L^*. The materials described above that given its contrast (c) and lightness (L^*) would not enforce

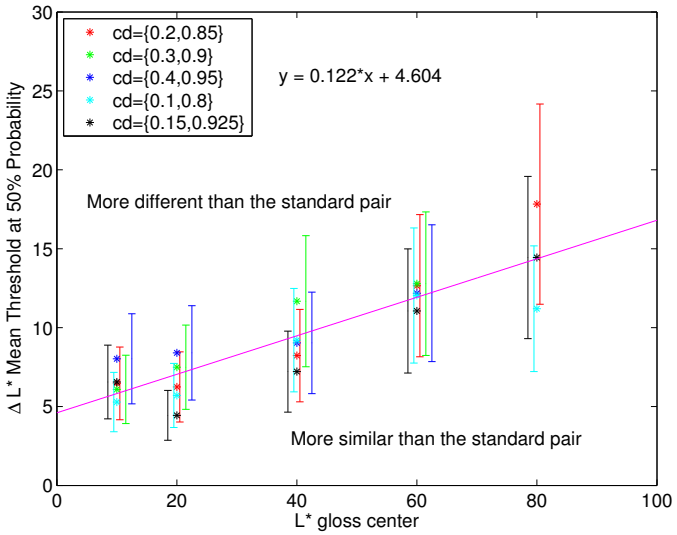


Figure 5.4: The observer's lightness discrimination was evaluated with gloss centers at different *cd* coordinates and lightness levels. The lightness difference threshold at 50% probability for each gloss center is shown as $*$.

energy conservation were not evaluated.

Results Experiment 1

For each test sample, the frequency of observers judging it to be closer to the gloss center than the standard pair was first calculated. Next, a gaussian function was fitted for each gloss center to the frequency of its test samples. The gaussian function was then used to obtain the lightness difference probability at 0.5, 0.75, and 0.25 from the gloss center, shown as $*$, lower, and upper error bars in Figure 5.4, respectively. The L^* perception was found to be independent of the *cd* position and the lightness differences were approximated with a linear function. Care should be taken when using the function outside the range studied $[10, 80]$.

The lighter the gloss center is, a greater change in lightness is needed for observers to perceive the test pair as being more different than the standard pair. Meaning that the observers lightness discrimination decreases as lightness increases. Thus, the perception of

lightness differences depends on the lightness of the samples.

The approximation obtained in Figure 5.4 can be used in the following form to compute the perceived lightness difference in the *Lcd* space:

$$f_L(L_i, L_j) = 0.05 \frac{|L_i - L_j|}{0.122\ L_i + 4.604} \tag{5.9}$$

where the 0.05 is the distance $D_{i,j}$ of the standard pair (which corresponds to the magenta line in Figure 5.4). Note that in this case $f_L(L_i, L_j) \neq f_L(L_j, L_i)$ and care must be taken to use the gloss center as L_i. Another option is to use the mean ($\frac{L_i+L_j}{2}$) instead of L_i in the denominator to compute the distance in the point located in the middle of the two samples, thus obtaining a symmetric function.

The revised distance function once lightness is added to the *Lcd* space is the following:

$$\Delta Lcd = \sqrt{[c_i - c_j]^2 + [1.78(d_i - d_j)]^2 + f_L(L_i, L_j)^2} \tag{5.10}$$

Discussion Experiment 1

In this gloss space L^* is used to define the material appearance of an object, and that appearance representation is later used to generate synthetic images. In this work, as in Pellacini et al. [2000] where the original space was defined, a LDR display is used, which requires the tone mapping of the images before being displayed. Tone mapping operators mostly compress the luminance of the scene, and this compression is probably being represented in the function obtained in Figure 5.4.

As future research, it would be interesting to study if there is a relation between the lightness difference function and the tone mapping operator used. At the same time, if the *Lcd* space is to be used to judge the perceived gloss differences between real objects, a similar experiment to the one performed in this work should be performed using either real

samples or a high dynamic range display, as it is known that limiting the image dynamic range does change the apparent gloss of surfaces depicted in images [Phillips et al., 2009]. Another reason why a function is required for L^* might be related to the same need as in the CIEDE2000 color difference equation.

5.2.2 Experiment 2: Gloss uniformity validation

The goal of the second experiment is to validate the uniformity of the *Lcd* space presented in Pellacini et al. [2000]. This validation will also allow the interpretation of the results of the third experiment, in which the projection presented in Section 5.1 will be evaluated. For this reason, the selection of the gloss centers for this experiment was guided by the needs of the third experiment.

To select the gloss centers for this experiment, the projection described in Section 5.1 was applied to all the 100 materials of the MERL database [Matusik et al., 2003]. Then, five of those projections were selected with the premise that they span most of the *Lcd* space and at the same time were far enough from the boundaries of the space to allow sampling around them. The projections of the materials depicted in Table 5.1 and Figure 5.5 were used as gloss centers for this experiment.

Only variations in the *cd* plane were considered for the test samples. The main reason behind this decision is because the lightness component of the projection defined in Section 5.1 is directly related to the 45:0 measurement, having a low uncertainty in comparison to the *cd* plane approximations. At the same time, the number of trials to perform is greatly reduced by only sampling two of the three dimensions of the *Lcd* space. To sample around each gloss center the Equation 2.27 was taken into account in order to have six perceptually equal steps in each dimension (step size $\Delta c = 0.033$ and $\Delta d = 0.033/1.78$) centered in the gloss center. Also, the gloss center itself was used as a test sample. As the test samples

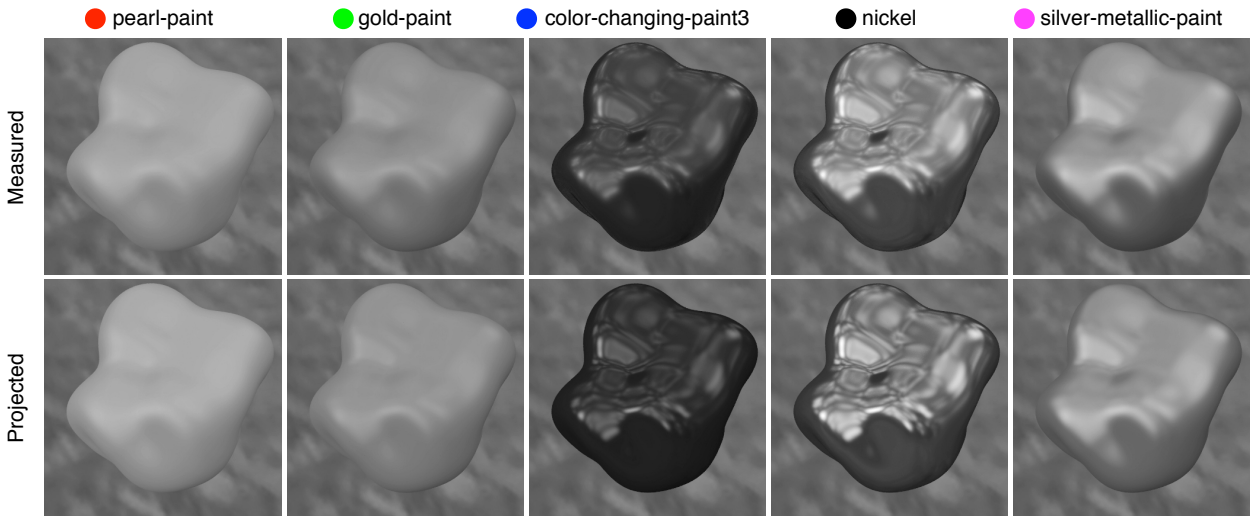


Figure 5.5: Measured materials used in the experiment (top row), and projections of those measured materials into the *Lcd* space (bottom row). The colors next to the material name will be used to reference a given material throughout the chapter.

Table 5.1: Projection of the measured materials into the *Lcd* space.

Material	L	c	d
pearl-paint	53.23	0.1333	0.7772
gold-paint	43.44	0.099	0.8415
color-changing-paint3	3.60	0.1567	0.9504
nickel	10.38	0.3969	0.9525
silver-metallic-paint	31.02	0.2878	0.8503

only varied in the *cd* plane, the standard pair for this experiment only varied in L^* and had $L^* = \{20, 15\}$ and $cd = \{0.2, 0.85\}$.

Results Experiment 2

The same idea of the MacAdam ellipses or color tolerances in CIELAB space will be used to evaluate the space uniformity. The gloss space will be uniform in a region if the difference of the standard pair to the test samples around the gloss center defines a circle, and the space will be non-uniform if an ellipse is obtained.

To compute each ellipse, the test samples that were selected by the observers to be more

similar than the standard pair with a frequency of 0.5 or higher were first selected. Then, a covariance matrix was defined with those test samples. Finally, the 95% confidence interval ellipse was obtained for the given covariance matrix and sample mean. As the sampling was selected to be uniform in the gloss space, the test samples were equally spaced. The ellipses obtained for each gloss center and the samples used to create the space in Pellacini et al. [2000] can be seen in Figure 5.6. In this case, the standard pair had a $\Delta Lcd = 0.036$ when using Equation 5.10.

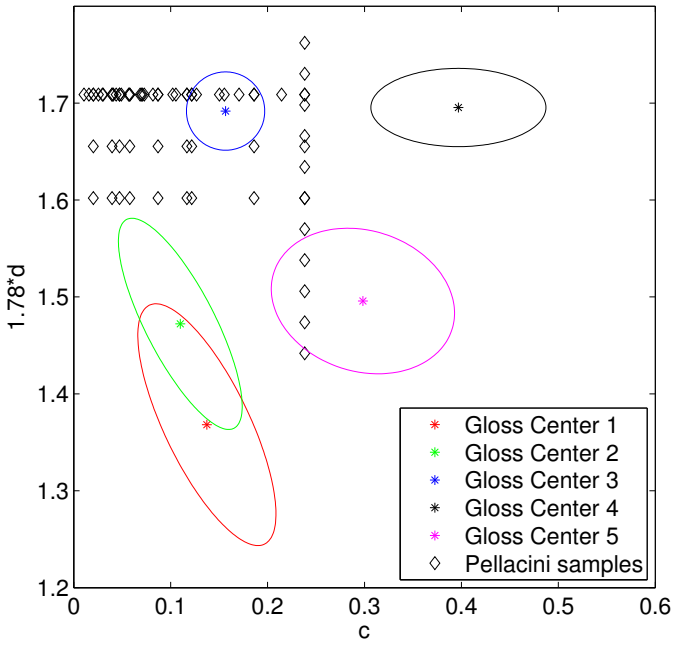


Figure 5.6: Each ellipse represents the set of materials that are perceived to be more similar to the gloss center ($*$) than the difference between the images of the standard pair. The ellipses found for the five gloss centers studied in the second experiment show the non-uniformity of the space outside of the samples used to create the gloss space in Pellacini et al. [2000].

Discussion Experiment 2

Circles would have been obtained instead of ellipses for each gloss center if the *cd* space was uniform. The ellipse determines the 95% confidence interval where 50% of the observers selected that the test pair was more similar than the standard pair.

Consequently the *Lcd* space defined in Pellacini et al. [2000] is non-uniform for the gloss centers tested. It's interesting to note that the original distance metric of the space seen in Equation 2.27 is only perceptually uniform around the top left gloss center studied, while being non-uniform for the rest of the gloss centers evaluated. In fact, the top left gloss center studied is the only one located between the samples used to define the gloss space in Pellacini et al. [2000] (see Figure 5.6). This finding agrees with the claim that the space only accurately predicts the appearance for materials in the range of the samples used to create the space, as stated by the authors [Pellacini et al., 2000].

More testing is indeed needed to better understand the non-uniformity of the space. From the small number of gloss centers evaluated in this experiment a trend can roughly be seen when looking at Figure 5.6. A compression in the c dimension might be needed with increasing c, as depicted by looking at the right-most ellipses. While a rotation and an additional compression in the d dimension for materials on the lower left quadrant might also be needed. It's also important to note that the different gloss centers have different lightness, as seen in Table 5.1, and lightness might also need to be considered when trying to improve the space uniformity. The observer discrimination seems to be correlated with lightness, as in the first experiment, as an increased ellipse elongation is observed with increased lightness.

Additional gloss centers throughout the space will be evaluated in Chapter 6 in order to validate the non-uniformity of the space obtained and to derive an improved gloss difference equation. A similar effect was seen with the CIELAB color space, where its euclidian color difference equation (CIE ΔE^*_{ab}) was refined over time and the CIE94 and CIEDE2000 were

later standardized.

5.2.3 Experiment 3: Projection Validation

The goal of the third experiment is to validate the projection of materials into the *Lcd* space defined in Section 5.1.

In this experiment, the same standard pair and test samples used in the second experiment were used. The only difference was that the gloss center used for this experiment was an image rendered using the measured data of the material instead of its projection into the *Lcd* space (see Figure 5.5).

The hypothesis is that if the projection performs well, similar ellipses will be obtained for the second and third experiments. The center of the ellipses obtained represents the location where the measured materials are perceived to be by the observers. Hence, the distance between the center of the ellipses of the second and third experiment will determine the projection distance error. In order to have the same perception of two materials, the distance between the center of the ellipses would need to be smaller than the JND in the *Lcd* space.

The JND in the *Lcd* space has already been discussed in the literature. First in Ferwerda et al. [2001], where a JND of around 0.03 was obtained with the data used to create the space. Later, in Vangorp et al. [2007] a detailed analysis of the discrimination threshold dependent on the object's shape and material was done. Where a threshold of 0.03 was found for the blob-like shape used in this study, and the JND was found to be dependent on the material used. The fact that the *Lcd* space is not uniform, as seen in Section 5.2.2, probably caused the material JND dependence reported in Vangorp et al. [2007]. Finally, in Vangorp and Dutré [2008] a JND of 0.0493 was considered to be near the material discrimination threshold.

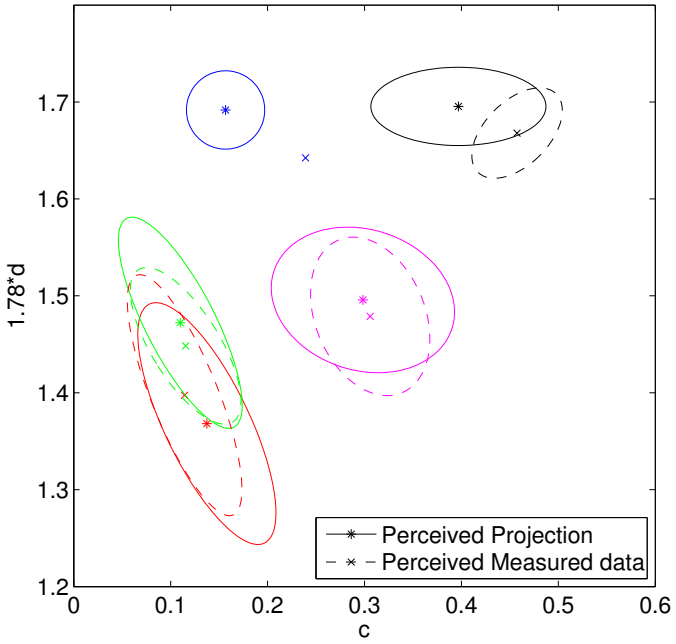


Figure 5.7: The perception of measured materials is represented as ellipses with dashed lines, and its projection into the *Lcd* space is represented as ellipses with continuous lines. Each material is shown in a different color. The projection presented is perceived to be near or below the JND threshold of the space for four of the five materials studied.

Results Experiment 3

The results of this experiment can be seen in Figure 5.7. The same procedure as described in Section 5.2.2 was used to compute the ellipses for this experiment.

The radius of an ellipse represents the perceptual distance of the standard pair, being $\Delta Lcd = 0.036$. Taking that into account, the distance between the perceived projection

Table 5.2: Perceptual distance between the measured data and its projection to the *Lcd* space.

Material	Projection distance error
● pearl-paint	0.011
● gold-paint	0.009
● color-changing-paint3	0.085
● nickel	0.034
● silver-metallic-paint	0.008

position and the perceived measured data in the space can be seen in Table 5.2. This distance is smaller than the JND of the space for the pearl-paint, gold-paint and silver-metallic-paint materials, close to the JND of the space for the nickel material, and larger than the JND for the color-changing-paint3 material.

A similar material perception is obtained for the pearl-paint and gold-paint using the projection, as highly coincident ellipses are obtained for those materials. No ellipse was obtained for the color-changing-paint3 material because only a single test sample was selected to be closer to the measured data for 50% or more of the observers.

Discussion Experiment 3

The accuracy of the projection defined in Section 5.1 has been validated in this experiment for five materials of the MERL database. The perceptual distance from the measured data to the projection obtained was found to be near or below the JND of the *Lcd* space for four of the five materials evaluated.

The projection distance errors found for the materials studied could be used to improve the projection defined in Section 5.1. However, care must be taken because the distance error found also include the limitations of the Ward BRDF model to accurately represent measured materials. The Ward BRDF model limitation to represent materials with broader tails [Löw et al., 2012] is probably the main reason behind the poor approximation obtained for the color-changing-paint3 material, as it can be seen in Figure 5.5.

5.3 Conclusions

A projection into the gloss space defined in Pellacini et al. [2000] was introduced in this chapter. This projection allows both measured data and analytical models to be represented in the space, thus allowing the computation of the perceptual distance of two materials.

This distance can be used to directly obtain an approximation of a measured material with the Ward BRDF model without using non-linear optimization, to guide an optimization procedure to approximate measured materials with other BRDF models, and for quality control applications.

Three different experiments were performed in this project. The first one was used to evaluate the lightness perception in relation to the contrast gloss and distinctness of image gloss dimensions of the space, and the lightness discrimination was found to decrease as lightness increased. The lightness perception was then modeled and included into the distance metric of the gloss space. The second experiment was used to evaluate the space perceptual uniformity, finding that the space was non-uniform for the positions evaluated and an improved distance metric is required. Finally, the last experiment was used to perceptually validate the projection to the gloss space introduced. The projection error was found to be near or below the JND threshold of the space for four of the five materials evaluated.

As the gloss space is a remapping of the Ward BRDF model, there are some measured materials that can't be accurately represented by the perceptual space, that being an inherent limitation of the presented projection. A metric to evaluate if a given material can be reasonably well approximated by the Ward model could be used to improve and better evaluate the projection defined in this chapter.

Material appearance measurements can be greatly simplified by using the projection defined. The use of simple instruments in conjunction with the projection described in this chapter to capture and represent material appearance will be evaluated in Chapter 7. Also, the non-uniformity of the space found in this chapter will be further studied in Chapter 6, and an improved gloss difference equation will be derived.

Chapter 6

Space uniformity improvement

The CIELAB color space and the color difference equations defined in it have been widely used for quality control applications and to characterize input and output devices taking into account human color perception. When considering gloss perception, perceptual attributes defined as ASTM features (distictness-of-image gloss, haze, bloom, etc.) are used for quality control applications, and perceptual gloss spaces are discussed in the computer graphics literature.

The perceptual gloss space defined in Pellacini et al. [2000] could be used for quality control applications to bring similar benefits as seen in color with the use of CIELAB. For example, the ability to set tolerances across multiple dimensions and the opportunity to have a single value to describe the perceptual distance between two materials. However, the space was found to be non-uniform outside the samples used to create it in Chapter 5.

In this chapter, the perceptual uniformity around eleven additional gloss centers in the gloss space was evaluated in a psychophysical study, which confirmed the non-uniformity of the space observed in Chapter 5, and an improved gloss distance measure of the space was derived.

Table 6.1: *Lcd* coordinates of the gloss centers studied in Chapter 5 (1-5) and Chapter 6 (6-16).

Gloss Center	L	c	d
1	53.23	0.1333	0.7772
2	43.44	0.099	0.8415
3	3.60	0.1567	0.9504
4	10.38	0.3969	0.9525
5	31.02	0.2878	0.8503
6	10.00	0.10	0.9
7	20.00	0.20	0.875
8	50.00	0.20	0.875
9	80.00	0.20	0.875
10	60.00	0.25	0.93
11	5.00	0.50	0.875
12	20.00	0.65	0.93
13	5.00	0.65	0.93
14	35.00	0.65	0.93
15	20.00	0.056	0.85
16	15.00	0.30	0.973

6.1 Experiment

In this experiment the space uniformity around additional gloss centers throughout the space was evaluated. This experiment was designed to better understand the non-uniformity of the space found in Chapter 5, to see if the same trends seen in Figure 5.6 are maintained throughout the space, and to aid the development of an improved gloss difference equation.

Eleven new gloss centers (6-16 in Table 6.1) were defined across the space with the goal to better understand the non-uniformity of the space, to evaluate materials with higher contrast and the ones closer to the boundaries of the space ($c = 0$ and $d = 1$), and to evaluate some cd positions at different L^* planes to analyze the relation of lightness with the other dimensions of the space.

The space uniformity around two different cd positions was evaluated at three L^* planes (gloss centers $(7, 8, 9)$ and $(12, 13, 14)$ in Table 6.1). To maximize the lightness range studied

the L^* planes evaluated for each cd position were different, as to enforce energy conservation materials with lower contrast (c) can reach higher lightness than high contrast materials.

The same experimental design described in Section 5.2 and the procedure of Section 5.2.2 was used in this study. The same number of observers, twenty-one, participated in this experiment and each performed a total of 399 judgements. The only difference was the sampling rate used for the gloss centers closer to the boundaries (15 and 16), being 4x8 (c x d) for the gloss center closer to $c = 0$ and 8x4 for the sample closer to $d = 1$. The same step size used in Section 5.2.2 was used.

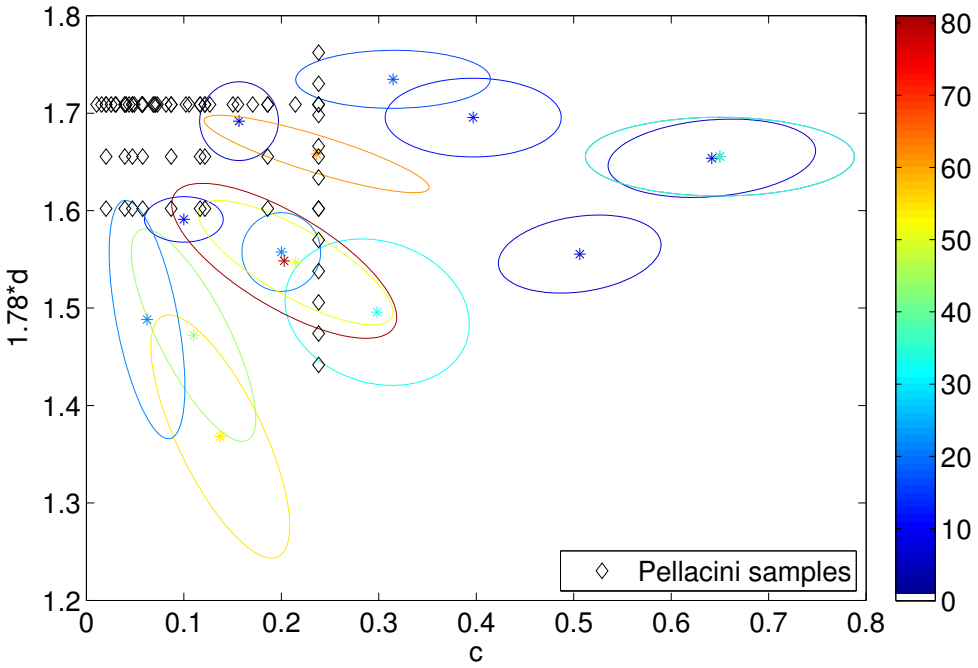


Figure 6.1: Ellipses obtained from the 16 gloss centers showing the space non-uniformity and the original samples used to create the space in Pellacini et al. [2000]. The ellipses are color-coded with its lightness plane (L^*).

Results

The same procedure used in Section 5.2.2 was used to evaluate the results from the psychophysical experiment, where the test samples that were selected by the observers to be more similar than the standard pair with a frequency of 0.5 or higher were first selected. Then, a covariance matrix was defined with those test samples. Finally, the 95% confidence interval ellipse was obtained for the given covariance matrix and sample mean. The ellipses obtained are shown in Figure 6.1 together with the original samples used to create the space in Pellacini et al. [2000].

An increased elongation in the *cd* plane is observed as L^* increases, as can be seen in the two *cd* positions evaluated at three lightness planes. Only two of the three different gloss centers at the $cd = \{0.65, 0.93\}$ position can be seen in Figure 6.1. This happens because the same test samples were selected by the observers with a frequency higher than 50% for the gloss centers 12 and 14, thus resulting with the same overlapping ellipse. At the same time, all the test samples available across the c dimension were selected for those two samples, meaning that a probably even larger horizontal elongation is perceived.

Discussion

The addition of eleven new gloss centers to the evaluation of the space uniformity allowed to confirm the trend seen in Section 5.2.2, where there seems to be an elongation in the c dimension with increasing c, an elongation on d as d decreases, and the orientation of the ellipses points out that there might exist a rotation point.

The increased elongation in the *cd* plane seen as L^* increases indicates that the observer discrimination decreases as L^* increases. The same effect was seen in Section 5.2.1, where the lightness discrimination also decreases with lightness. However, a more complex interaction between L^* and the *cd* dimensions is apparent from the results obtained, as the elongation

is dependent on the cd position of the sample. For the samples in the $cd = \{0.2, 0.875\}$ position, the space goes from being uniform to show an increased elongation as L^* increases, and for the samples in the $cd = \{0.65, 0.93\}$ position the space is already stretched at low lightness and it stretches even more as L^* increases. It's interesting to note that ellipse size of the gloss center with lowest lightness ($L^* = 5$) with high contrast ($cd = \{0.65, 0.93\}$) has almost the same size as the ellipse with mid contrast ($L^* = 50$) in the low contrast region ($cd = \{0.2, 0.875\}$).

A further evaluation is probably needed to better understand and be able to model the cd dimensions dependence on L^*. At the same time, it would also be interesting to further validate the uniformity of the region where the samples used to create the space in Pellacini et al. [2000] are located and in addition evaluate that region at multiple lightness planes.

6.2 Improved gloss difference equation

In this section an improved gloss difference equation will be derived using the results from the psychophysical experiment described above.

The standardized residual sum of squares (STRESS) metric, presented in García et al. [2007] and commonly used in the color science community to evaluate color difference equations, will be used to evaluate the performance of existing and developed gloss difference equations. STRESS values measure the deviation between visual differences (ΔV) and numerical differences (e.g. ΔE^*_{ab}), while at the same time allowing to make statistical inferences of two different equations.

For each gloss center, the distance from the perceived location of the gloss center in the space ($*$ in Figure 6.1) to each point that defines the ellipse represents an equal visual difference, as the ellipse represents the materials that would have the same distance to the gloss center as the distance between the samples in the standard pair. To compute STRESS,

the data for all gloss centers is used and compared to a given numerical difference equation applied to each pair of samples in the space. Because of the resemblance to its expressions and for easy understanding, reference to the color difference equations will be used throughout the following explanation.

First, the euclidian distance metric in the *Lcd* space (ΔE^*_{cd}) was computed for the gloss centers and ellipses obtained in this work, giving a STRESS value of 44.27. Then, the STRESS value using the gloss difference equation derived by Pellacini et al. [2000] (Equation 2.27), was computed and a STRESS of 39.76 was obtained. Lower STRESS values indicate a lower deviation between visual differences and numerical differences, meaning that an improvement was obtained by scaling the d dimension by 1.78.

Space Modeling

The development of the improved gloss difference equation was guided by the goal to model the elongation in the c and d dimensions, to model the rotation observed in the ellipses, and with the assumption that the space is uniform in the region where the samples used to create the space are located.

The starting point of the modeling process was Equation 2.27, as it is assumed that it provides a perceptual uniform space where the samples used in Pellacini et al. [2000] are located.

The use of cylindrical coordinates was a natural evolution in color difference equations, where the distance from neutral and the rotation angle is used to describe chroma and hue in CIELAB, respectively. The same approach was used in this work to improve the gloss difference equation of the space. In this case, the rotation point is not as clearly defined, neither conceptually nor as a point in the space. Still, the orientation of most ellipses seems directed towards a point near the upper left limit of the space, $cd = \{0, 1.78\}$. As a

conceptual meaning for the distance and rotation from the rotation point in terms of gloss was not found, the same nomenclature, chroma and hue, as in CIELAB will be used for clarity.

Two optimization processes were done to first find the rotation point of the space, and then to model the increased ellipse elongation seen in Figure 6.1 as the gloss centers are farther a part from the rotation point. For the first optimization process, the rotation point (cd) and the function for chroma that will be described later were non-linearly optimized to minimize the STRESS value. The sixteen ellipses found experimentally were used and the rotation point obtained was $Rot_{cd} = \{0.004, 1.686\}$. The rotation point represents the point in the space that better aligns with the long axis of all the ellipses.

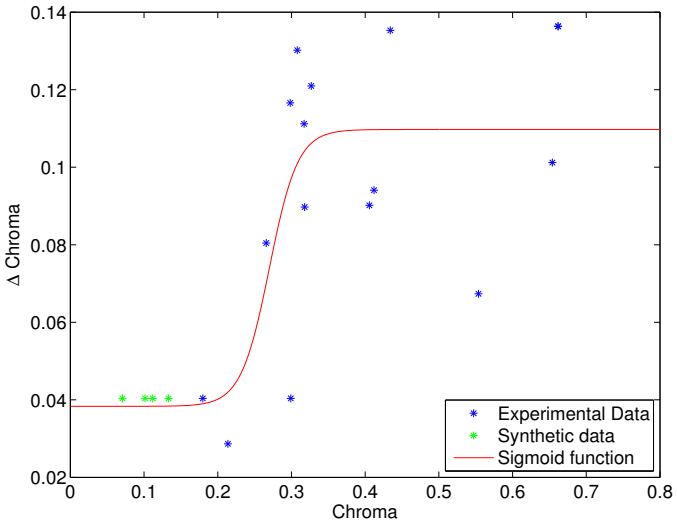


Figure 6.2: $Chroma$ vs $\Delta Chroma$ for the 16 gloss centers obtained in this study, the four synthetic circles added to enforce the uniformity in the area where the Pellacini et al. [2000] samples were located, and the sigmoid used to approximate the data.

Once the rotation point was known, a second optimization process was performed to add a weighting function to chroma. Chroma defines the distance from the rotation point and directly relates to the increased ellipse elongation as the gloss centers are farther a part from the rotation point. In this case, four synthetic circles equal in size to the top left circle

seen in Figure 6.1 were placed in the region where the Pellacini et al. [2000] samples are located in order to enforce the space uniformity in that area. That decision lead to the good performance of a sigmoid function, as can be seen in Figure 6.2. Chroma is the distance from the rotation point to the gloss center (see Equation 6.1), and ΔChroma is the radius of the ellipse in the direction of the rotation point from the center of the ellipse. The four additional circles added and the sigmoid function that approximates both the experimental and synthetic data are also shown in Figure 6.2.

For the second optimization process, the parameters of the sigmoid function were non-linearly optimized to minimize the STRESS value. In this case, both the sixteen ellipses found experimentally and the four synthetic circles were used during the optimization process. The sigmoid parameters obtained were the following: $p = \{0.87, 2.66, 0.15, 3140.9\}$. Finally, the first parameter (p_1) was set to 1 in order to enforce the region uniformity where the Pellacini et al. [2000] samples are located. A STRESS value of 22.96 was obtained when the sixteen ellipses found experimentally were evaluated, which was found to be a statistically significant improvement when compared to the ΔE^*_{cd} and the $D_{i,j}$ gloss difference equation from Pellacini et al. [2000].

To compute the gloss difference equation the distance from the rotation point for each sample (Chroma) is first computed:

$$C_1 = \sqrt{[c_1 - Rot_c]^2 + [(1.78d_1) - Rot_d]^2} \tag{6.1}$$

$$C_2 = \sqrt{[c_2 - Rot_c]^2 + [(1.78d_2) - Rot_d]^2} \tag{6.2}$$

where the rotation point $Rot_{cd} = \{0.004, 1.686\}$, and cd are the coordinates of the two samples evaluated. Next, the chroma and hue differences between those samples are computed

as in ΔE^*_{94}:

$$\Delta c = c_1 - c_2 \tag{6.3}$$

$$\Delta d = d_1 - d_2 \tag{6.4}$$

$$\Delta C_{cd} = C_1 - C_2 \tag{6.5}$$

$$\Delta H_{cd} = \sqrt{\Delta c^2 + \Delta d^2 - \Delta C_{cd}^2} \tag{6.6}$$

The sigmoid function used in Figure 6.2 is the following one:

$$S_C = p_1 + \frac{p_2 - p_1}{1 + 10^{[(p_3 - C_1)p_4]}} \tag{6.7}$$

where $p = \{1, 2.66, 0.15, 3140.9\}$ are the sigmoid parameters. Finally, the gloss difference equation is computed:

$$\Delta G^*_{cd}(Lcd_1, Lcd_2) = \sqrt{f_L(L_1, L_2)^2 + \left(\frac{\Delta C_{cd}}{S_C}\right)^2 + \Delta H_{cd}^2} \tag{6.8}$$

where Lcd_1 and Lcd_2 are the coordinates of the two samples, S_C is used to weight ΔC_{cd}, and Equation 5.9 is used to compute f_L.

The approximation of the experimental results with the derived gloss difference equation can be seen in Figure 6.3, where the resulting ellipses from the experiments are shown in dashed black lines, and the ellipses obtained with the ΔG^*_{cd} are shown in red. In order to compare the size and orientation, the same scaling used to match the experimental data of the top left circle with the gloss difference equation was used for all the other ellipses. As C and H are independent from L^*, the results seen for the gloss centers studied at three different lightness planes are not represented, and the same ellipse is obtained in each cd position.

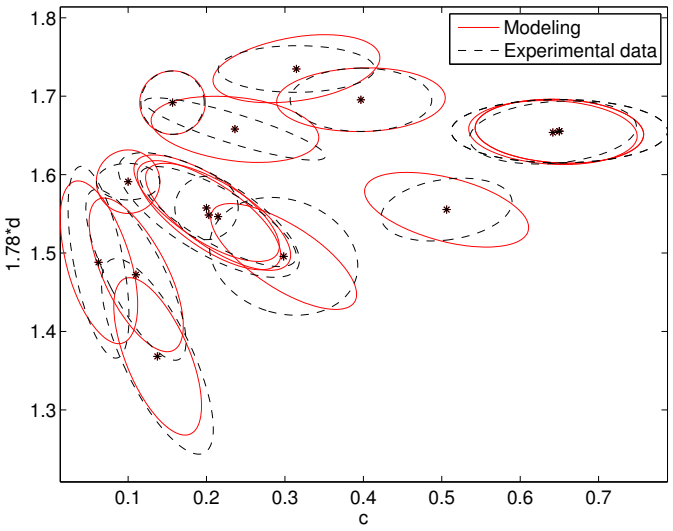


Figure 6.3: Ellipses obtained in the psychophysical studies are visualized with dashed black lines, and results obtained using the gloss difference equation ΔG^*_{cd} developed are shown in red.

Figure 6.4 shows the behavior of the ΔG^*_{cd} at different locations across the *cd* plane at any L^* level. The space is uniform in the region where the samples used to create the space in Pellacini et al. [2000] were located, while outside this region the ellipses are elongated in respect to the rotation point.

6.3 Conclusions

In this chapter, the non-uniformity of the gloss space found in Chapter 5 was validated by evaluating the perceptual uniformity around eleven additional gloss centers in the gloss space. Then, the space non-uniformity was modeled and an improved gloss difference equation was defined. A more detailed analysis is needed to better understand how L^* effects the space uniformity of the *cd* plane, and for that reason this interaction was not included in the gloss difference equation presented. As a future work it will also be interesting to validate the uniformity of the gloss space at different lightness planes in the region where the samples

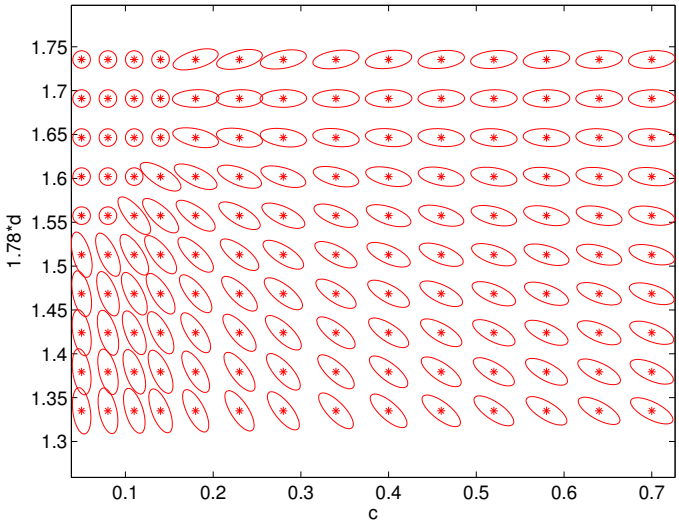


Figure 6.4: Ellipses displayed across the *cd* plane in order to show the function behavior across the space. Note that the *cd* plane shown is independent of lightness.

used to create the space are located.

Finally, the perception of color in relation to the gloss space would be an interesting avenue of future work to obtain an overall material perception space including both perceptual attributes, color and gloss.

6.4 Evaluation of the media dependent gloss perception study

In this section, the results of the cross-media gloss perception study of Chapter 3 will be evaluated in the *Lcd* space. The projection presented in Chapter 5 will be used to represent the materials into the *Lcd* space, and the improved gloss difference presented in this chapter will be used to evaluate the results obtained.

In Chapter 3, three different gloss discrimination experiments were performed. In the first one, the gloss matching performance with real samples was evaluated. The second

experiment evaluated the matching ability when using synthetic images as representations of real objects. Finally, the last experiment evaluated the discrimination ability on the cross-media situation, where real objects were matched with synthetic representations.

The material approximations used to render the images of the experiment conducted in Chapter 3 were used in this section. The high accuracy measurements from the ELDIM EZContrast could also have been used to evaluate the results of the first experiment, where the matching experiment between real objects was evaluated, but for consistency across the evaluation of the three experiments only the material approximations using the Ashikhmin-Shirley were projected into the *Lcd* space and used in this section.

After the materials were projected to the gloss space, for each experiment and for each reference sample the mean *Lcd* coordinate across the set of samples selected by the observers to be the best match to the reference sample was computed. This mean *Lcd* coordinate represents where the reference sample was perceived in the *Lcd* space for a given experiment.

Finally, the gloss difference equation derived in Section 6.2 was used to perceptually evaluate the differences between the location of the sample references in the *Lcd* space, and the location of the perceived reference sample obtained from the psychophysical experiment.

Results

The projection of the material approximations using the Ashikhmin-Shirley BRDF model obtained in Chapter 3 to the *Lcd* space can be seen in Figure 6.5. The eight reference samples that the observers had to match in each of the studies are shown in red.

The projection of the results obtained in Chapter 3 into the *Lcd* space can be seen in Figure 6.6, and the arrows show the difference and the direction from the reference sample to the mean perceived location of the reference sample in the space. The mean, minimum, maximum, standard deviation, and 90^{th} percentile of the ΔG^*_{cd} (Equation 6.8) across the

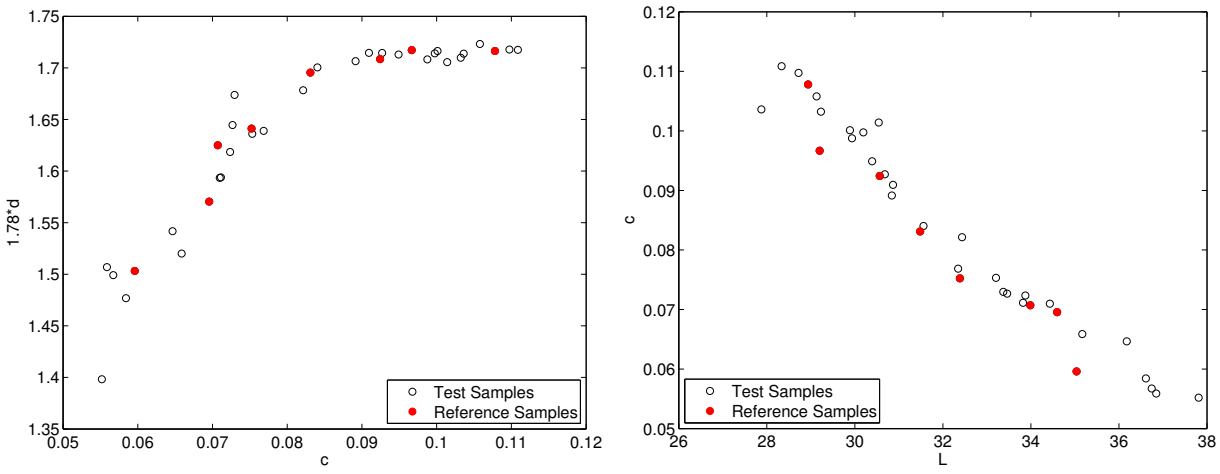


Figure 6.5: Approximations of the 36 samples with the Ashikhmin-Shirley BRDF model used in Chapter 3 projected into the *Lcd* space, *cd* coordinates (left) and *Lc* coordinates (right). Reference samples used for the matching experiment are shown in red.

Table 6.2: ΔG^*_{cd} across the reference samples obtained from the results of Chapter 3.

	Real vs. Real	Display vs. Display	Display vs. Real
mean	0.5697	0.5592	0.5184
min	0.3646	0.1668	0.1336
max	0.8531	0.9802	1.143
std	0.1839	0.3136	0.3062
90^{th} percentile	0.8363	0.9439	1.0105

eight reference samples for each experiment can be seen in Table 6.2.

Discussion

The 36 gloss samples were created in Chapter 3 by applying a varying amount of Digital Matte (0-300%), a varnish that decreases the gloss, on a high gloss substrate. As the amount of digital matte is increased, contrast gloss decreases linearly with lightness, as can be seen in Figure 6.5 right. On the other hand, the application of Digital Matte up to 90% only decreases contrast, while keeping a constant distinctness of image gloss (specular lobe width), and samples with more than 90% of Digital Matte decrease in both dimensions, contrast gloss

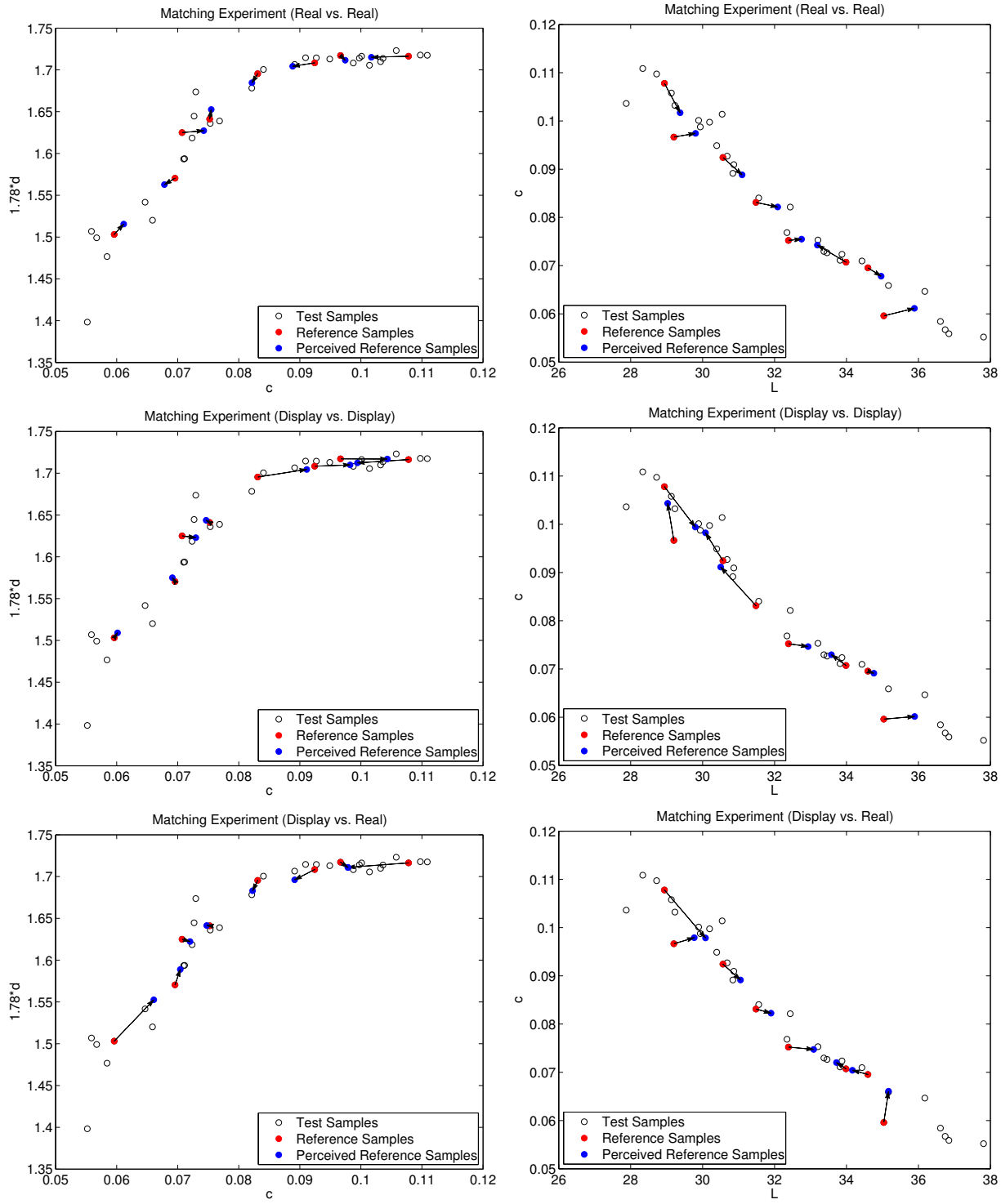


Figure 6.6: Projection of the results obtained in Chapter 3 for the three experiments into the *Lcd* space, *cd* coordinates (left) and *Lc* coordinates (right). The arrows show the difference and direction from the reference sample to the mean perceived location of the reference sample in the space.

and distinctness of image gloss (see Figure 6.5 left).

In Chapter 3, the reference samples were visually selected to be equally spaced in terms of perceived gloss. This process could have been simplified and improved with the use of the projection described in Chapter 5, the *Lcd* space presented in Pellacini et al. [2000], and the improved gloss difference equation presented in this chapter. The design of future studies could benefit from this process, as it enables the computation of accurate perceptual distances between samples and the understanding of their distribution in the gloss space before performing any psychophysical study.

The observer deviations from the reference samples observed in Figures 3.6, 3.7, and 3.8 are also perceived with the projected approximations of the materials and the experiment results in Figure 6.6.

For almost all the reference samples of the first experiment (Real vs. Real) the samples selected by the observers as a match had a lower contrast and higher lightness, following the sample distribution across those two dimensions. Caution must be taken to not model the effect seen relying only on the data obtained, as the trend found is directly influenced by the distribution of the samples used in the space. The reference sample should be sampled as in Section 5.2.2 or in the three dimensions to accurately capture the observers perception.

A low accuracy is observed for the high gloss samples in the second experiment (Display vs. Display), where a higher contrast gloss is perceived for three of the four reference samples in that region. Small differences in the magnitude of the specular peak alone for high gloss samples are probably difficult to perceive on a display due to the human perception, where the sensitivity observed for high lightness is lower than for low lightness, as commonly exploited in gamma encoding. The lack of binocular cues, as stated in Chapter 3, can also contribute to this lower accuracy to discriminate between high gloss samples.

The same trend was obtained in the first experiment (Real vs. Real) and in third ex-

periment (cross-media scenario, Display vs. Real), where the observers underestimated the contrast gloss and overestimated the lightness of the samples. The accuracy for high gloss samples, except for the glossiest sample, is higher than the results obtained for the second experiment, where only rendered images are used, and might indicate that the presence of the real object allows to better judge the gloss perception of objects shown on a display.

The mean gloss difference perceived across all the reference samples is slightly higher than the upper bound of the space JND considered in the literature, being from 0.3 to 0.5 [Ferwerda et al., 2001; Vangorp et al., 2007; Vangorp and Dutré, 2008], which might indicate that the JND of the space is closer to the upper bound than the lower. The minimum gloss difference found for a reference sample was smaller for the second and third experiments, where a display was used, and it might be related to the higher observer precision in areas of high discrimination obtained by using the keyboard to easily switch and compare different samples, instead of placing the physical samples one at a time in the light booth. The maximum and 90^{th} percentile of the gloss difference across reference samples increase as more visual constrains were added to the experiment design, as also seen in Figure 3.9.

Additional cross-media gloss perception studies should be done with samples across the space to further validate and model the effects seen in this work.

Chapter 7

Abridged gloss measurement

Man-made objects have a geometry that is easy to model with current CAD applications and 3D scanning techniques to be used in virtual environments. However, the capture of material appearance is a more challenging task. Goniospectrophotometers can be used to measure how the light is reflected from a material at different directions, but the time consuming measurement process and expensive nature of those devices limits its use.

The generation of realistic synthetic images relies on an accurate representation of the material appearance of objects, and improvements on its capture and representation would improve the realism of previews of to be printed 3D objects, regular 2D printing taking into account the appearance of the substrate and inks used, and the accurate representation of real materials in computer generated images in general.

Lab grade and commercial devices are available to perform BRDF measurements. Some examples of lab grade solutions are: The goniophotometer used at Cornell University to perform BRDF measurements [Foo, 1997], the technique used in Matusik et al. [2003] to perform high angular resolution measurements of the 100 isotropic materials in the MERL database, the NIST reference goniospectrophotometer [Obein et al., 2005], or the goniometer capable of performing BRDF and BTDF (T, for transmission) at Lawrence Berkeley National

Lab (with a similar design as Apian-Bennewitz [2010]). Those setups provide high accuracy measurements and in some cases also high angular resolution measurements. However, those setups are expensive to build and maintain, and the long measurement time required makes them not scalable.

Commercial measurement devices have also been developed for specific applications and can be used to perform BRDF measurements, e.g. the Murakami GCMS-10x goniospectrophotometer, the Eldim EZContrast, and the Radiant Zemax Imaging Sphere. The time required to measure BRDFs depends on the amount of information to be captured, but they are usually faster than the lab grade setups outlined above. However, those devices are also expensive limiting its widespread use.

Several approaches have been proposed in the computer graphics literature to speed up the measurement process and represent real materials in synthetic images. In Westlund and Meyer [2001], a virtual goniometer able to simulate a gloss meter and a haze meter was derived. Allowing to calculate the gloss and haze values for each set of parameters of a given analytical BRDF model. Then, a look-up-table was created to relate a set of measurements with a set of BRDF parameters to use at rendering time. Other approaches that use cell phone cameras to obtain a compelling representation of real materials were presented in Ren et al. [2011] and Wang et al. [2011].

In this chapter a cost-effective, fast, and scalable solution to capture the material appearance is presented. The main idea behind this work is the use of simpler devices, commonly used for quality control applications, and to combine the measurements to represent the material appearance of an object. Devices used for quality control applications are used in this work as they are fast, as multiple samples need to be measured in a short period of time and to constrain the measurements to important perceptual properties.

The material information to capture was split into two main attributes: color and gloss.

A spectrophotometer was used to capture the color of a material, and the raw data of a linear sensor used in a DOI-Haze meter was used to obtain BRDF measurements, thus capturing the gloss appearance. Those measurements were later used to approximate the parameters of analytical BRDF models.

The technique presented was evaluated by comparing its results with the high accuracy measurements of a goniospectrophotometer, and the approximations obtained when the approximation process starts with the high accuracy measurements. The comparison between the measurements and approximations were also evaluated using the perceptual gloss space defined in Pellacini et al. [2000] and the gloss difference equation defined in Chapter 6.

7.1 Measurement Technique

The main idea behind the measurement technique presented in this chapter is to split the information necessary to represent the material appearance into two different attributes: color and gloss. The same separation is used to represent BRDFs with analytical models, where a diffuse lobe is commonly used to represent the color of an object, and a specular lobe is used to represent the gloss appearance. The technique presented would generalize for any isotropic material, but the monochromatic capture of the specular lobe by the device used in this work limits the current applicability of this technique to dielectrics. The color of metals is known to be represented in the specular lobe and thus the technique can not be applied for metals.

The diffuse reflectance of a material is acquired using the 45:0 measurement geometry (illumination:measurement directions), this geometry is commonly used in colorimetry to avoid capturing the specular component. The X-Rite i1 spectrophotometer was used in this work to capture the spectral reflectance of a material, and a PTFE created from pressed teflon powder was used as standard. The CIE 2° standard observer and illuminant D65 were

used to obtain CIE XYZ values from the measured spectral reflectance data.

The requirements for gloss measurements are the following: high angular resolution, high dynamic range, and the measurement of multiple incident directions. The high angular resolution is required to correctly capture the width of the specular lobe. The high dynamic range is required to be able to capture materials from diffuse that have almost negligible specular lobes to high gloss. Finally, the ability to measure at different incident directions enables the measurement of the Fresnel effect behavior, which models the increase in reflectance when the incident direction goes towards grazing angles.

The Rhopoint IQ is a DOI-Gloss-Haze meter used for quality control applications. Its design is similar to a 20/60/85 gloss meter, but a linear sensor of 512 pixels covering $\pm 7.25^\circ$ around the mirror direction at 20° is used instead of a single diode to provide an angular resolution of 0.028°. The Rhopoint IQ reports Specular Gloss at the three different geometries, and by using the linear sensor the DOI, Haze, RSPEC (average gloss at the peak of the specular lobe), and a goniometric curve are reported for the 20° geometry. The gloss attributes reported by the device are relative attributes computed against the black glass standard used during calibration.

By default, the goniometric curve is reported in gloss units, and is expected to only be used as an aid to better understand the material angular reflectance of the sample measured. Access to the RAW sensor data was provided to me by the manufacturer, thus enabling the use of the device to perform BRDF measurements. The measurement time of the Rhopoint IQ is about 5 seconds.

The diffuse reflectance of a material is set directly to the 45:0 measurement, and the parameters of the analytical BRDF model used can be non-linearly optimized to approximate the BRDF data measured with the Rhopoint IQ.

The Rhopoint IQ characteristics limit the applicability of this technique to non-metallic

and non-goniochromatic materials, and as only a single incident direction is captured with the linear sensor the increased reflectance towards grazing angles (Fresnel effect) behavior is not captured. Thus, the fact that some materials might have a different behavior when the incident direction goes towards grazing angles can not be measured.

The following equation is used to compute the amount of light being reflected by a material using the Rhopoint IQ:

$$i_{sample} = \frac{RawMeasurement - DarkCurrent}{IntegrationTime} \tag{7.1}$$

Then, the BRDF is obtained by using the following equation:

$$f_{sample} = \frac{i_{sample}}{i_{PTFE}\,\pi} \tag{7.2}$$

where i_{PTFE} is the measurement with the Rhopoint IQ of a perfect reflectance diffuser created with pressed teflon powder, and π is the normalization factor.

Commonly, BRDFs are approximated using the following equation:

$$f_{BRDF} = \rho_d + \rho_s\ specularLobe \tag{7.3}$$

where ρ_d approximates the diffuse reflectance and is set directly to the 45:0 measurement, and the parameters of the *specularLobe* analytical BRDF model and ρ_s can be non-linearly optimized to approximate the gloss measurements.

7.2 Goniospectrophotometer Comparison

In this section, the accuracy of the technique presented above will be evaluated by measuring a set of samples with a goniospectrophotometer and using them as ground-truth. The set of

36 printed samples with varying gloss used in Chapter 3 was used to evaluate the technique presented in this chapter.

The Murakami GCMS-10x goniospectrophotometer, which measures the spectral reflectance factor as a function of incident and detection angles, was used as a reference instrument. The 20° incident direction was measured, as in the Rhopoint IQ, and the outgoing directions measured covered the range $-70°$ to 75° in θ, having a higher sampling rate near the mirror direction (20°). The directions farther than 20° from the mirror direction were sampled every 5°, directions from 10° to 20° were sampled every 1°, and directions from the mirror direction up to 10° were sampled every 0.5°. No samples were measured between $-40°$ and 0° due to the occlusion between the light source and the sensor. The measurement time with the Murakami was about 10 minutes for each sample. The CIE 2° standard observer and illuminant D65 were used to obtain CIE XYZ for each of the angular geometries, and CIE Y was used to compare to the measurement obtained with the Rhopoint IQ, which is the integration over its sensor sensitivity. Its important to note that the comparison will be between CIE Y and the integration over the Rhopoint IQ's sensor sensitivity, due the lack of photometric information for the Rhopoint IQ.

The Rhopoint IQ measurements of the 36 samples and the PTFE used for calibration, after applying Equation 7.1, can be seen in Figure 7.1. Five measurements with and without replacement (measuring at slightly different positions of the sample) were performed on the samples to evaluate the precision of the instrument. The Rhopoint IQ showed a good precision for the glossy samples, but the results were not consistent for the PTFE measurements. Differences in the integration time lead to large noise differences in the PTFE measurements, due to the auto exposure algorithm of the device expecting glossy materials. Several measurements were taken and the one with the longest integration time, having less noise, was finally used.

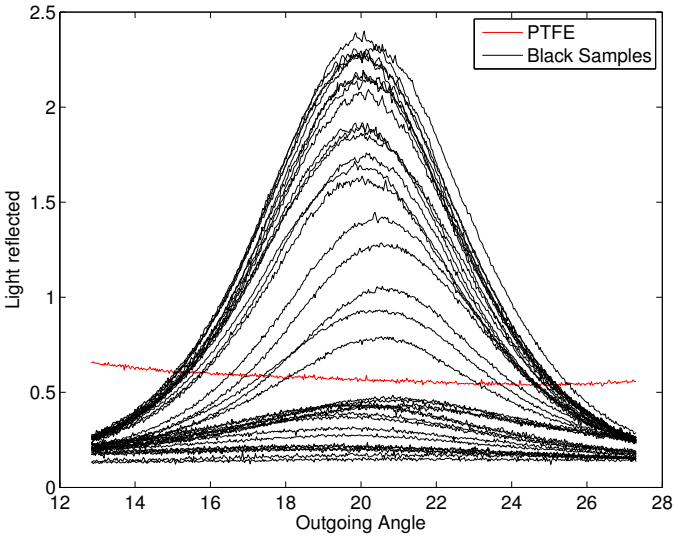


Figure 7.1: Rhopoint IQ measurements of a PTFE and the 36 black samples with different gloss levels used for evaluation.

Ideally, the PTFE measurement would be represented as a horizontal line in Figure 7.1, as light is reflected from a PTFE equally in all directions. The use of each geometry, the mean, and the minimum of the PTFE measurement were evaluated as i_{PTFE} in Equation 7.2 to better understand the the implications of the PTFE measurement obtained when comparing it to the measurements of the goniospectrophotometer. The use of $min(i_{PTFE})$ in Equation 7.2 was finally used as it provided the best approximation of the goniospectrophotometer CIE Y measurements, while the other approaches had a more pronounced underestimation of the specular peaks.

Figure 7.2 shows the measurements of the technique described in the previous section and the CIE Y measurements of the Murakami. A good approximation of the Murakami measurements is obtained with the Rhopoint IQ across the different gloss samples, except for a small underestimation of the specular peak of high gloss samples (Samples 1-12), and the small shift of the specular peak observed for mid gloss samples (Samples 16-24).

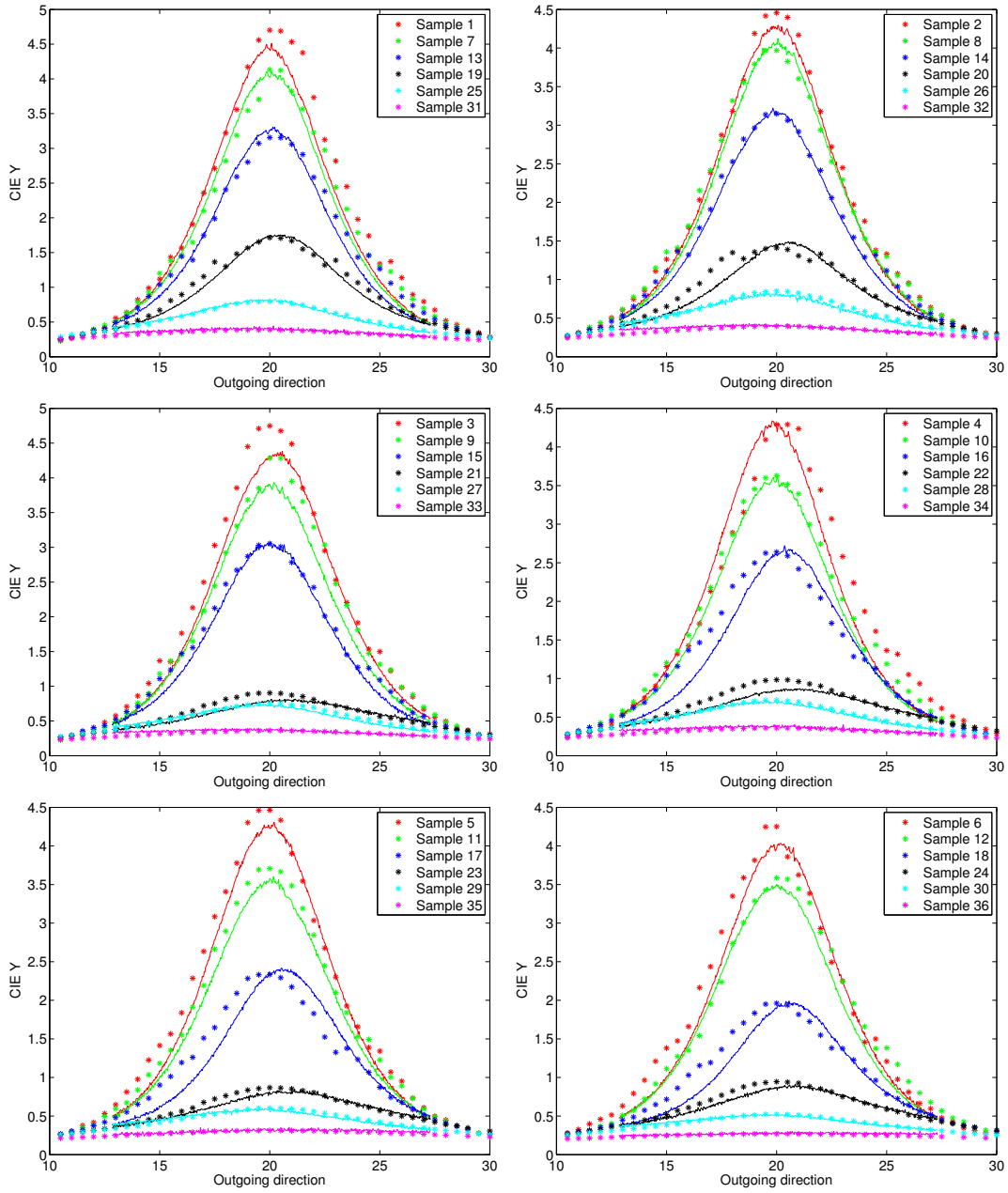


Figure 7.2: Murakami GCMS-10x measurements (∗) and Rhopoint IQ measurements (continuous lines) of the 36 black samples with varying gloss levels. The same color is used to represent the same material measured with both devices.

Discussion

In overall, the Rhopoint IQ was able to correctly approximate the measurements of the high accuracy goniospectrophotometer. The main advantage of using the Rhopoint IQ,

and the reason why it is presented in this project as a solution is the dramatic reduction in measurement time and cost achieved by using the Rhopoint IQ when compared to the options currently available.

The initial design considerations of the Rhopoint IQ lead to the precision and accuracy limitations seen when measuring diffuse materials, and for that reason the measurement with the longest integration time was selected and the minimum of the PTFE measurement was used for calibration. Modifications to the auto exposure algorithm of the device to take into account a broader range of materials to be measured would be able to handle the integration time variability observed.

The design differences between the Rhopoint IQ and the Murakami are probably the main cause of the underestimation of the specular peak seen by the Rhopoint IQ. The Murakami measures each geometry (incident:outgoing direction) separately, and with the use of several neutral density filters is able to measure samples spanning a high dynamic range difference with high accuracy. However, the Rhopoint IQ performs all the measurements in a single exposure, at the expense of its accuracy and high dynamic range. Another difference is the fact that the Murakami CIE Y is compared to the Rhopoint IQ sensor integration, which is not photometric.

The Murakami is able to obtain accurate measurements of the samples evaluated in this project, which have a medium to low gloss (see Figure 3.3) and a relatively broad specular lobe. The Murakami sensor integrates a large sample area, which can smooth out data of mirror-like surfaces and result in broader BRDF measurements, and for that reason it is not well suited to measure materials with sharp specular lobes. On the other hand, the high angular resolution of the Rhopoint IQ sensor would provide a higher accuracy than the Murakami in respect to the width of the specular lobe for materials with sharp specular lobes (e.g. mirror-like surfaces). HDR techniques could be applied to the Rhopoint IQ in

the future to improve its dynamic range by combining multiple exposures, thus allowing accurate measurements throughout the different measurement geometries.

7.3 Measurement Approximation

Analytical BRDF models are commonly used for rendering, as they present a compact representation and importance sampling can be used to speed-up the rendering process. For that reason, measured data is commonly approximated with analytical models to achieve those benefits while representing the measured material. In this section, the approximations obtained with both measurement techniques, the goniospectrophotometer and the technique presented in this chapter, will be compared.

In the approximation procedure, the parameters of a given analytical BRDF model are non-linearly optimized in order to minimize a given BRDF error metric. The selection of analytical BRDF models and error metrics was studied in Chapter 4 and in the literature [Ngan et al., 2005; Löw et al., 2012]. To select the BRDF model best suited for the set of 36 materials studied in this project the Ward [Ward, 1992], Ashikhmin-Shirley [Ashikhmin and Shirley, 2000], and Smooth Surface BRDF [Löw et al., 2012] models were evaluated, and the cube root metric (Equation 4.3) was used during the optimization process, as compressive metrics were found to obtain a better perceptual approximation than regular RMS metrics in Chapter 4, in Matusik [2003], and in Löw et al. [2012].

A Lambertian lobe was used to represent the diffuse reflectance (p_d in Equation 7.3) and it was set to the measurements taken at a 45:0 geometry, being commonly used in colorimetry to remove specular reflections. The non-linear optimization of the diffuse reflectance was also tested, leading to an overestimation of the diffuse reflectance for the high gloss materials used, as more weight was given to the specular lobe, and a more accurate approximation of the diffuse reflectance for low gloss materials. For simplicity and consistency, the 45:0 geometry

was used across all the materials.

The Smooth Surface BRDF model (Equation 2.16) was used for this study as it gave the best approximation for the measured data obtained in the previous section. The Ward and Ashikhmin-Shirley BRDF models have an underlying gaussian distribution, which do not approximate the shape of the measured data well, specially the tail of specular lobes [Löw et al., 2012]. If the specular lobe is approximated with a gaussian function, the near-diffuse region is constrained by the gaussian distribution to be narrower than the measured data. An example of this can be seen in the top left plot of in Figure 7.6, where the measured data is approximated with the Ward BRDF model.

The goodness of the approximations will be evaluated by inspecting the BRDF shape of the measurements and the approximations for both devices, and by comparing the rendered images using the approximations obtained with both devices.

The Murakami and Rhopoint IQ measurements and approximations of six representative materials of the ones studied can be seen in Figure 7.3. To avoid overemphasizing the specular peak magnitude, the plots are compressed with the cube root. The approximations from both devices correctly reassemble their measured data, and similar approximations using both devices are obtained for the Samples 1, 8, 19, and 34, while more substantial difference is obtained for the Samples 29 and 35. The approximation with the Rhopoint IQ is more narrow for the Sample 29 and broader for the Sample 35 than the Murakami approximations.

Images rendered using the approximations obtained with the Murakami and Rhopoint IQ measurements for six representative materials can be seen in Figure 7.4. In order to maximize the material discrimination in the rendered images a blob-like shape and the Eucalyptus Grove light probe from Paul Debevec were used. This geometry gave the best material discrimination accuracy in Vangorp et al. [2007] and this environment map was

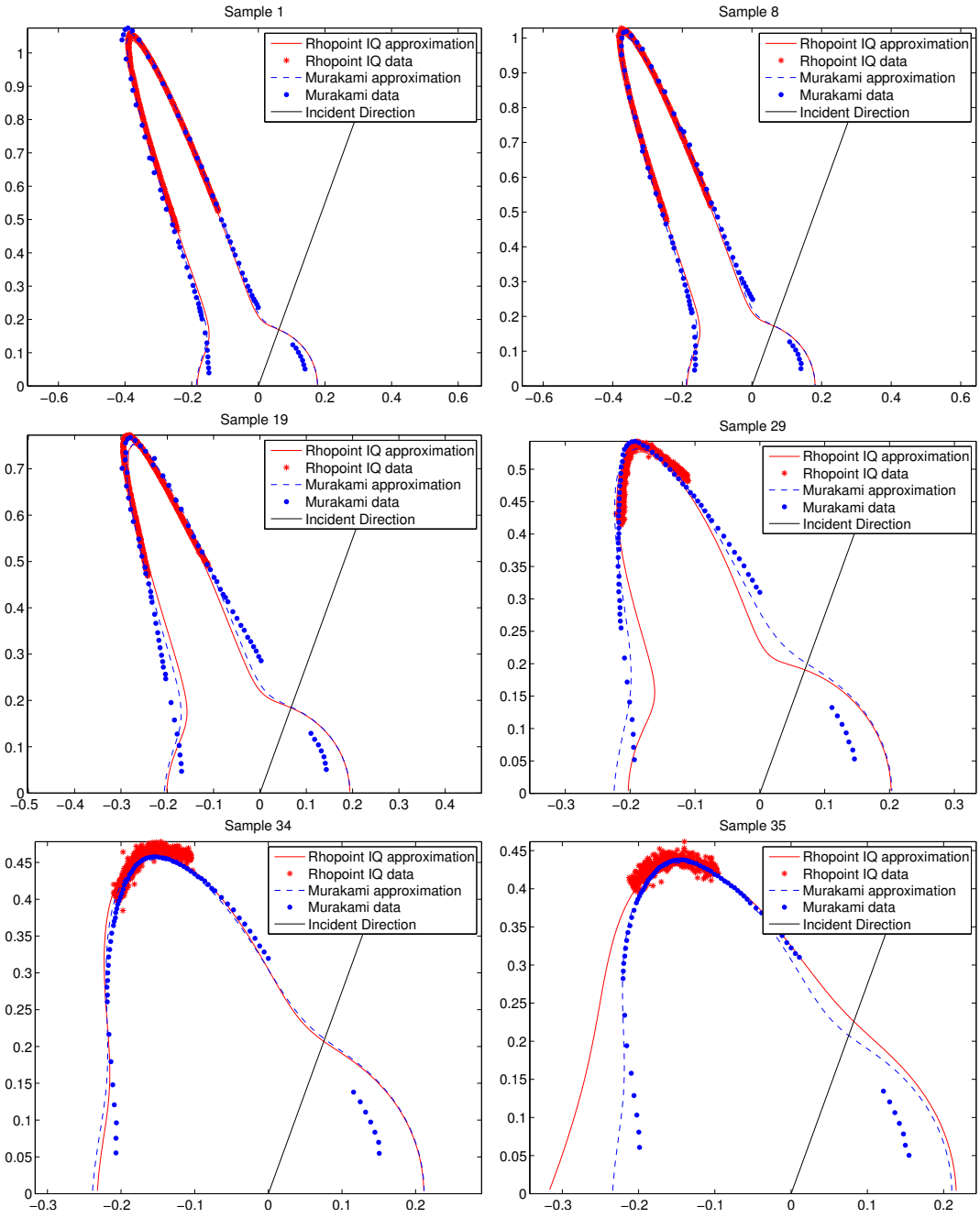


Figure 7.3: Approximation of the Rhopoint IQ and Murakami measurements using the Smooth Surface BRDF model [Löw et al., 2012] and the cube root error metric for 6 samples. To avoid overemphasizing the specular peak magnitude, the plots are compressed with the cube root.

found to be the environment map with real world statistics providing the best material

discrimination in Fleming et al. [2003], respectively. Visually, the same perception is obtained from the materials approximated with both devices, except for a small perceptual difference for samples 29 and 35. The Rhopoint IQ approximation of Sample 29 appears to have sharper reflections, while the opposite happens to Sample 35, where the reflections are broader and a more diffuse appearance of the material is perceived than the approximation obtained with the Murakami measurements. Those material appearance changes seen on the images correlate with the findings observed in the BRDF shapes seen in Figure 7.3.

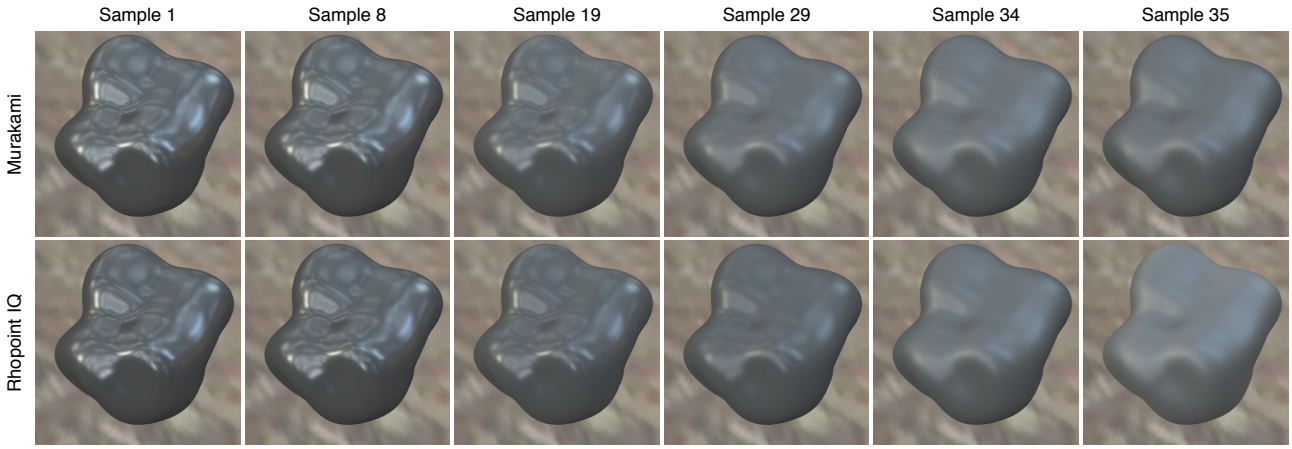


Figure 7.4: Images of the approximations obtained from the Rhopoint IQ and Murakami measurements with the Smooth Surface BRDF model [Löw et al., 2012].

Discussion

The most distinct features that can be seen in Figure 7.3 are the sampling density and the range of angles measured with each device. The Rhopoint IQ provides a high angular resolution of a limited angular range, while the Murakami's flexibility enabled the measurement of a much broader range of angles.

For the high gloss samples studied, the Rhopoint IQ is able to capture most of the specular lobe with the exception of the near-diffuse region. However, the Rhopoint IQ

measurements of low gloss samples only capture the almost flat specular peak, without providing any information about the width of the specular lobe. This limitation in angular range results in the lack of accuracy of the approximation for the samples 29 and 35. As there are no Rhopoint IQ measurements constraining the width of the specular lobe for samples 29 and 35 the optimization process does not consider that factor on its minimization and different outcomes might occur, having either a narrower or a broader approximation as seen in samples 29 and 35, respectively. On the other hand, the Murakami measurements were spaced to capture the different regions of the specular lobe: the peak, the width, and the near diffuse region, correctly capturing each feature.

The selection of the Smooth Surface BRDF model after testing different analytical BRDF models using both sets of measurements also influenced the goodness of the approximations obtained with the Rhopoint IQ across the materials. For example, if the analytical BRDF model selection would have been done without the Murakami measurements, the Ward or Ashikhmin-Shirley models might had been considered, as the limitation in the near-diffuse region of those models would not be apparent. For that reason, in spite of the goodness of the Rhopoint IQ approximations obtained, the BRDF model needs to be carefully selected to represent the type of material being measured. This limitation could be overcome by modifying the lens design of the Rhopoint IQ in order to increase its angular range, or by having a previous knowledge about the type of materials that are going to be measured and select the analytic BRDF model accordingly.

Another important factor that is not evaluated in this project is the modeling of the increasing reflectance when the incident direction goes towards grazing angles. This effect, known as the Fresnel effect, is perceived when the viewing direction moves away from the normal direction, increasing its reflection as the direction moves farther away from the normal direction. The degree for which a material increases in reflectance is material dependent

and must be measured to correctly characterize it. As the Rhopoint IQ only measures a single incident direction the same was done for the Murakami measurements, in spite of the flexibility to measure multiple incident directions with the latter device. Thus, being able to asses the accuracy of the measurements and the approximations obtained with the Rhopoint IQ. Indeed, the Frensel effect of the approximations of the Rhopoint IQ or the Murakami shown in this project might not be accurate as only the measured data of a single incident direction was used to drive the optimization. To overcome this limitation, another geometry could be added to the Rhopoint IQ (e.g. 60°) to capture some information about the Fresnel effect and be able to model it during the approximation process.

Another interesting avenue of future work would be the addition of the measured data as rendered images, to compare the goodness of the approximations with both devices in comparison to the reference material. A higher angular resolution than the one captured with the Murakami, and the need to sample out of plane would be needed to render the measured data directly. For example, as done in Matusik et al. [2003] for the materials of the MERL database.

The technique presented captures the color information of the diffuse lobe with a spectrophotometer. This simplification is valid for dielectric materials, where its color is defined in the diffuse reflection and the specular lobe is considered to be achromatic. However, this simplification is not valid for metals, where the color information of the material is contained in the specular lobe. As the Rhopoint IQ has a monochromatic sensor this technique is not well suited to measure metallic materials as its color would not be captured. The addition of a RGB filter array in the linear sensor would allow to capture the colorimetric information of the specular lobe while at the same time preserving a reasonable angular resolution, and combining this with the measurement of two incident directions would also enable to capture goniochromatic materials.

7.4 Projection to a gloss space

In the previous section, the BRDF shape of the approximations was compared to the measured data, and rendered images of the approximations were used to evaluate the technique presented in comparison to a high accuracy device. In this section, the projection presented in Chapter 5 will be used to represent the measurements and approximations obtained in Sections 7.2 and 7.3 into the gloss space defined in Pellacini et al. [2000], and the improved gloss difference equation derived in Chapter 6 will be used to evaluate the perceptual difference between the measurements from the Rhopoint IQ and the Murakami, and to compare it to the perceptual difference between the approximations obtained with both devices.

The incident direction measured on the Rhopoint IQ is 20°, which differs from the 30° used to compute the ASTM standards for the projection. The relation between gloss ASTM standards and d was recreated for 20° and the measurements with the Murakami at 20° were projected into the Lcd space with the updated relation. Also, the 36 materials of this study were measured with the incident direction at 30° with the Murakami and projected to the Lcd space. Both sets of measurements in the Lcd space were compared and the ΔG^*_{cd} was computed. A mean ΔG^*_{cd} of 0.014 was found between the measurements at 30° and the measurements at 20°, with the only difference of the measurements at 30° having a slightly higher contrast than the measurements at 20°. This change in contrast is due to the increased reflection as the incident direction goes towards grazing angles (Fresnel effect) commonly seen on materials. This difference is smaller than the JND of the space, considered by the literature to be between 0.3 and 0.5 [Ferwerda et al., 2001; Vangorp et al., 2007; Vangorp and Dutré, 2008], meaning that the materials are equally perceived in the space, and with this assumption the 20° projection was used to evaluate the measurements and approximations obtained in the previous sections.

The relation between the gloss ASTM standards and d is used in the projection to

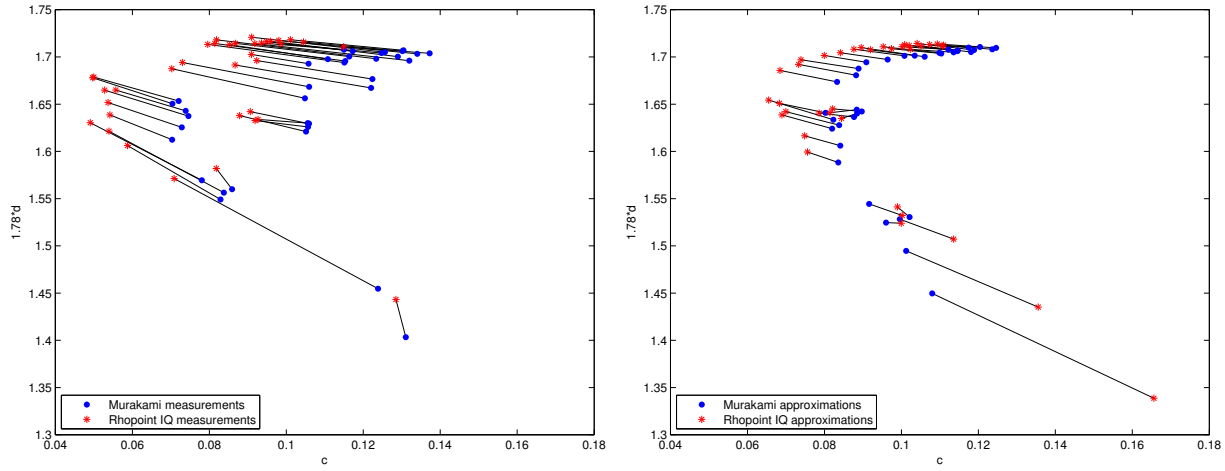


Figure 7.5: Rhopoint IQ and Murakami measurements projected into the *Lcd* gloss space (left), and Rhopoint IQ and Murakami approximations using the Smooth Surface BRDF model projected into the *Lcd* gloss space (right). The black lines connect the same samples across datasets.

determine the width of the specular lobe in the *Lcd* space. Due the lack of measurements for the features located at 10° and 15° with the Rhopoint IQ, and additional feature within the angular range of the Rhopoint IQ was added at 7° to increase the accuracy of the relation between ASTM standards and *d*.

Figure 7.5 left shows the projection of the Murakami and Rhopoint IQ measurements into the *Lcd* space, and Figure 7.5 right shows the projection of the approximations obtained in the previous section using the Smooth Surface BRDF model to the *Lcd* space. The black lines connect the same samples across datasets.

The gloss difference equation defined in Chapter 6 (Equation 6.8) was computed between the samples projected to the *Lcd* space for each device in each case to understand the magnitude of the distance perceived in the *Lcd* space in respect to the JND of the space. The mean ΔG^*_{cd} between the Murakami and the Rhopoint IQ measurements projected into the *Lcd* space is 0.025, and the mean ΔG^*_{cd} between the Murakami and Rhopoint IQ approximations is 0.013. Thus, the difference between the measurements performed with the

Murakami and Rhopoint IQ and the approximations using the Smooth Surface BRDF model are below the JND of the space, with the exception of the measurements of the Sample 35 ($\Delta G^*_{cd} = 0.053$), and the approximation of the Sample 36 ($\Delta G^*_{cd} = 0.055$).

As the gloss space from Pellacini et al. [2000] is a reparameterization of the Ward BRDF model, the position in the *Lcd* space corresponds to a material described with the Ward BRDF model. Thus, the projection for arbitrary BRDFs presented in Chapter 5 can also be used to approximate measured materials without the need to perform non-linear optimization. Figure 7.6 shows the BRDF shapes for the measured data and its projection to the *Lcd* space for the Murakami and Rhopoint IQ measurements, and Figure 7.7 shows the images rendered with the projection of the measured data to the *Lcd* space for both devices.

Discussion

A clear trend in Figure 7.5 left is seen between the measurements on the Murakami and the Rhopoint IQ in the contrast dimension, where the Rhopoint IQ measurements have a lower contrast than the Murakami. Also, the Rhopoint IQ measurements have a slightly higher d, meaning that a sharper specular lobe is obtained. The same trends are observed with the approximations obtained using the Smooth Surface BRDF model (Figure 7.5 right), with the exception of the low gloss materials (small d) for which a higher contrast (c) is obtained for the Rhopoint IQ approximations. The differences are more pronounced on both dimensions for materials with broader specular lobes (small d in Figure 7.5). Larger differences are seen for low gloss samples due the low accuracy of the Rhopoint IQ to measure the width of the specular lobes of low gloss materials, which also impacts the accuracy in the c dimension.

The approximations with the Smooth Surface BRDF model projected into the *Lcd* space (Figure 7.5 right) follow a smooth variation from sample to sample, while the measurements projected into the *Lcd* space (Figure 7.5 left) are more scattered. This effect is probably

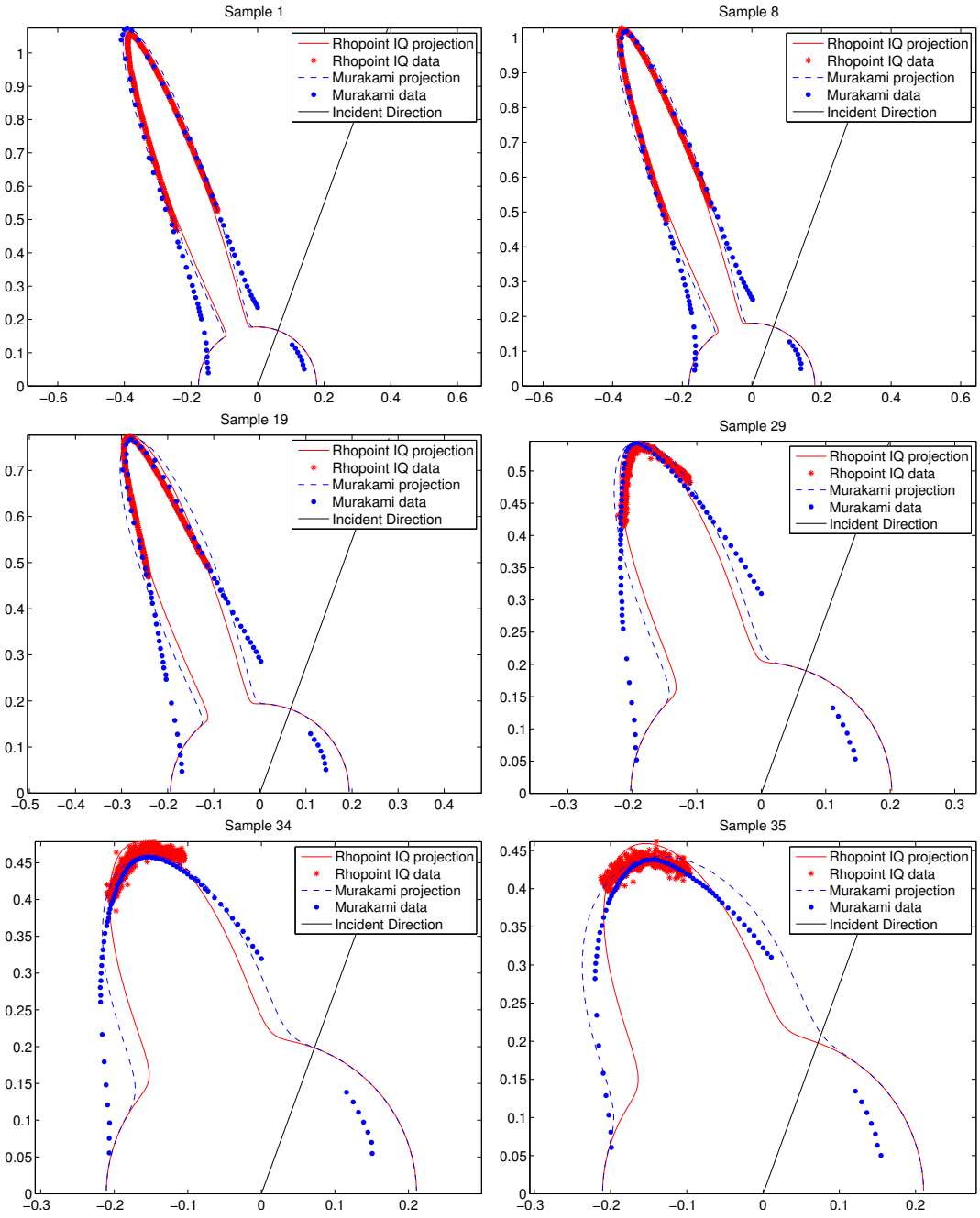


Figure 7.6: Projections into the *Lcd* gloss space of the Rhopoint IQ and Murakami measurements using the projection described in Chapter 5. To avoid overemphasizing the specular peak magnitude, the plots are compressed with the cube root.

caused by the measurement precision of the instruments and a small surface non-uniformity

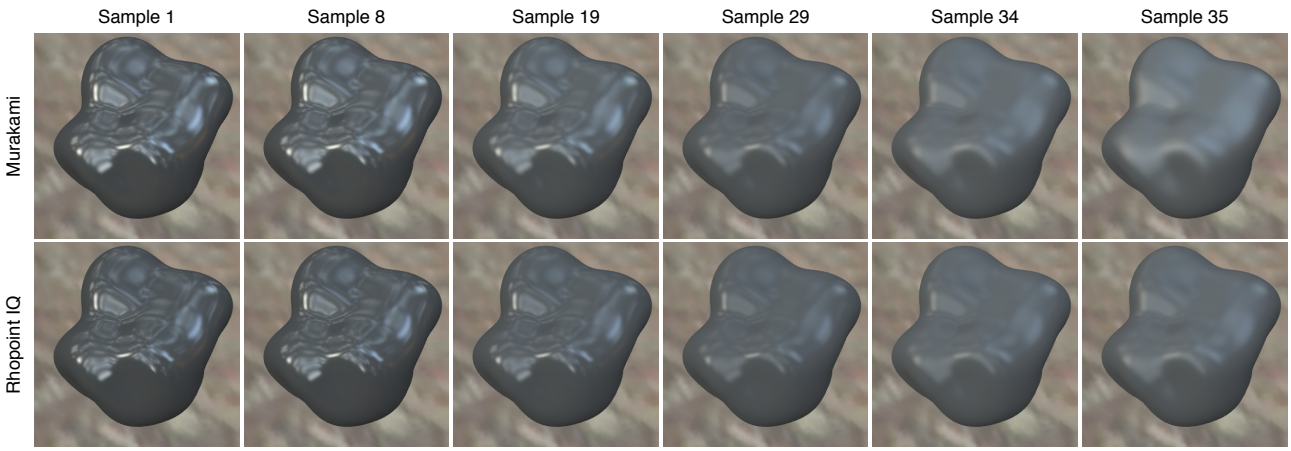


Figure 7.7: Images of the Rhopoint IQ and Murakami measurements projected into the *Lcd* gloss space.

of the samples, which is then smoothed in the approximation step, leading to smaller gloss differences between the approximations done with both devices for each sample.

Figure 7.6 shows the BRDF shapes of the measured data and its projections into the *Lcd* space. The projection into the *Lcd* space of the measurements for both devices achieves a good approximation of the specular peak and the upper part of the specular lobe. However, the near diffuse region is approximated as being narrower than the measurements for both the Murakami and the Rhopoint IQ instruments. The narrow specular lobe is obtained because of the projection to the *Lcd* space, that gives higher weight to the width of the specular lobe at half of the magnitude of the specular lobe, and the magnitude of the specular peak. At the same time, it is a limitation of the underlying BRDF model used in the space, the Ward BRDF model, which can not accurately represent both the upper part of the specular lobe and the near-diffuse region at the same time [Löw et al., 2012]. It's also interesting to note that the Murakami measurements projected into the *Lcd* space always have a broader specular lobe than the Rhopoint IQ, closely reassembling the measurements. The fact that the Murakami had the ASTM features at 10° and 15° improved the approximation of the d parameter, which directly relates to the width of the specular lobe. This finding correlates

with the results found in Section 7.3, where the wide range of angles measured with the Murakami provided a higher accuracy that could not be achieved with the Rhopoint IQ due the limited angular range captured.

The rendered images of the measurements for both devices projected to the *Lcd* space, seen in Figure 7.7, show a sharper representation of the materials than the approximations obtained in Section 7.3. The main reason for this difference is the underlying analytical BRDF model used in the *Lcd* space, the Ward BRDF model, for which its gaussian distribution limits the representation of the near-diffuse region [Löw et al., 2012]. Also, the differences between the measurements of the Rhopoint IQ and the Murakami are more noticeable than the rendered images of its approximations using the Smooth Surface BRDF model (see Figure 7.4).

Materials projected into the *Lcd* space, and the differences perceived in it, are represented with the Ward BRDF model. The limitation of this model to represent the near-diffuse region of the specular lobe limits the approximation of certain materials, as seen in Figures 7.6 and 7.7. This limitation implies that the gloss difference derived in Chapter 6 (ΔG^*_{cd}) does not take into account the differences seen in the near-diffuse region of measured materials, as the near-diffuse region of the projections of materials into the *Lcd* space might not be accurately represented. An interesting avenue of future work would be the addition of a near-diffuse feature into the gloss difference equation of the space. The recreation of the gloss space using a BRDF model capable of representing the behavior of the near-diffuse region, like the Smooth Surface BRDF model, would be another option to overcome this limitation.

7.5 Conclusions

To conclude, a novel cost-effective, fast, and scalable solution to measure material appearance is presented in this project. The main contribution is the separation of the measurement of different appearance properties of materials, color and gloss, while using simpler devices to perform each measurement. A good approximation was obtained when comparing the new technique to a goniospectrophotometer, except for a small underestimation of the peak of the specular lobe of high gloss materials and the limitation to capture the specular lobe width of broad specular lobes found in low gloss materials.

The validation of the proposed technique using a broader range of materials and other high accuracy devices would be desired in the future, in order to further validate the generalization of the solution proposed in this project.

The use of the Rhopoint IQ enables the technique presented in this work, but at the same time its design constrains the usage of the presented technique to non-metallic materials, due to the achromatic capture of the specular lobe, and the Fresnel effect is not captured as only a single incident angle is measured.

A promising avenue of future work is the creation of a measurement device that builds upon the initial Rhopoint IQ design and addresses the limitations found in this chapter. Having a RGB filter array in the linear sensor would allow to capture colorimetric information of the specular lobe while still preserving reasonable angular resolution. A change in the lens design would allow to capture beyond the current $\pm 7.25°$ from the mirror direction, thus being able to correctly capture the specular lobe width of low gloss materials. Finally, the addition of a second linear array at another measurement geometry would allow to (1) capture the Fresnel effect behavior and (2) capture goniochromatic effects when combined with the RGB filter. The improvement of the sensor's dynamic range would also lead to an overall accuracy improvement.

Chapter 8

Conclusions and Future Work

This chapter summarizes the contributions of this thesis. It outlines the different projects related to perceptual modeling and reproduction of gloss discussed throughout the thesis and it provides new directions for future work.

8.1 Conclusions

The media dependent gloss perception was first studied in this thesis. Similar trends were observed when judging the gloss of physical objects and in the cross-media scenario, where the observers underestimated the contrast gloss and overestimated the lightness of the samples. When only a display was used to perform judgements, a low contrast gloss accuracy for medium gloss samples was found. Still, additional cross-media gloss perception studies should be done with samples distributed across the space to further validate and model the effects seen in this work.

The visual fidelity obtained by different error metrics when approximating measured materials with analytical models was perceptually evaluated in this thesis. Compressive metrics were found to obtain a better perceptual representation of measured materials when

used as objective function in the optimization process. Similar compressive functions are used in CIELAB to represent lightness and in tone mapping operators to compress HDR imagery to be displayed on low dynamic range displays. Compressive metrics give a higher weight to the near-diffuse region of the specular lobe, which was also found to be a key feature in gloss reproduction in Löw et al. [2012].

A novel projection for arbitrary BRDFs into the perceptual gloss space defined in Pellacini et al. [2000] was defined. This projection improves the usefulness of the gloss space and it could bring similar benefits to gloss reproduction as the use of CIELAB did for color. The space could be used to represent the gamut of analytic BRDF models and measurement devices, and the perceptual error metric defined in it could be used to drive optimization procedures, to name a few. At the same time, the space was found to be non-uniform outside the region where the samples used to create it are located and an improved gloss difference equation was derived.

A cost-effective, fast, and scalable solution to capture material appearance was also presented in this thesis, where simpler devices commonly used for quality control applications were used to measure different properties of materials, color and gloss, and its measurements were used to represent real materials in synthetic images. This technique could broaden the adoption of measured data to represent real materials in synthetic imagery, by avoiding the time consuming and complex adjustment of parameters to obtain the desired material appearance and at the same time improving its perceptual accuracy.

In summary, the novel contributions of this dissertation are:

- The evaluation of the media dependence gloss perception.
- The perceptual evaluation of the visual fidelity obtained with different error metrics when used to approximate measured data with analytical BRDF models.

- A projection for arbitrary BRDFs to the gloss space defined in Pellacini et al. [2000].
- An improved gloss difference equation that models the space non-uniformity found outside the samples used to create it in Pellacini et al. [2000].
- A cost-effective, fast, and scalable solution to capture material appearance.

8.2 Extensions and Future Work

The work presented in this thesis encourages more research in similar directions, which can broaden and improve the applicability of perceptual gloss reproduction.

8.2.1 Validation of the gloss space for real materials

The projection presented in Chapter 5 and the perceptual gloss space defined in Pellacini et al. [2000] has been validated for low dynamic range displays. To enable the use of the space with physical objects directly, and thus be able to use it for quality control applications the space and the projection defined in this thesis would need to be validated with real materials.

The perception across the space with real materials could be evaluated by performing similar experiments to the ones conducted in Chapters 5 and 6. In this thesis, the analytical definition of the space with the Ward BRDF model simplified the creation of samples around the gloss centers studied to evaluate the space uniformity.

The recent work in the computer graphics literature to create real objects with arbitrary BRDFs could be used to define gloss centers across the space and samples around them. For example, a computer-controlled machine tool was used to mill a surface to obtain a specific BRDF in Weyrich et al. [2009], different BRDFs were obtained by printing with specialty inks in Matusik et al. [2009] and in Dong et al. [2012], and the materials with arbitrary subsurface scattering were manufactured in Matusik et al. [2009].

8.2.2 Material appearance space combining color and gloss

The study of gloss perception for achromatic materials has been studied in this thesis, except for the evaluation of error metrics done in Chapter 4 were color was also present. The combination of color and gloss would enable to better describe the overall material appearance difference between two materials, bringing clear benefits to quality control applications and material reproduction in synthetic imagery.

Again, similar experiments as the ones seen in Chapters 5 and 6 could be conducted to study several gloss centers across the space, defined with different lightness, hue, and chroma. This would help to improve the understanding between the relation of color and gloss, and would help to improve the gloss difference metric relation of lightness with the *cd* dimensions, as seen in Chapter 6.

8.2.3 Improved hardware design for abridged BRDF capture

The use of the Rhopoint IQ enables the abridged gloss measurement technique presented in this thesis, but at the same time its design constrains the usage of the technique to non-metallic and non-goniochromatic materials, and the accuracy of the Fresnel behavior approximation.

A promising avenue of future work is the creation of a measurement device that builds upon the initial Rhopoint IQ design and addresses the limitations found in this thesis. Having a RGB filter array in the linear sensor would allow to capture colorimetric information of the specular lobe while still preserving reasonable angular resolution. A change in the lens design would allow to capture beyond the current $\pm 7.25^{\circ}$ from the mirror direction, thus being able to correctly capture the specular lobe width of low gloss materials. Finally, the addition of a second linear array at another measurement geometry would allow to (1) capture the Fresnel effect behavior and (2) capture goniochromatic effects when combined with the RGB filter.

The improvement of the sensor's dynamic range and the use of HDR techniques would also lead to an overall accuracy improvement.

References

Agarwal, S., Wills, J., Cayton, L., Lanckriet, G., Kriegman, D., and Belongie, S. (2007). Generalized non-metric multidimensional scaling. In *AISTATS*, San Juan, Puerto Rico.

Apian-Bennewitz, P. (2010). New scanning gonio-photometer for extended brtf measurements. In *Proc. SPIE*, volume 7792.

Ashikhmin, M. and Shirley, P. (2000). An anisotropic phong BRDF model. *J. Graph. Tools*, 5:25–32.

Blinn, J. F. (1977). Models of light reflection for computer synthesized pictures. In *SIGGRAPH '77: Proceedings of the 4th Annual Conference on Computer Graphics and Interactive Techniques*, pages 192–198, New York, NY, USA. ACM.

CIE (2006). A Framework for the Measurement of Visual Appearance. Technical report, CIE 175:2006.

Cook, R. L. and Torrance, K. E. (1982). A reflectance model for computer graphics. *ACM Trans. Graph.*, 1:7–24.

Dana, K., Van-Ginneken, B., Nayar, S., and Koenderink, J. (Jan 1999a). CUReT: Columbia-Utrecht Reflectance and Texture Database. [web page] http://www1.cs.columbia.edu/CAVE//software/curet/.

Dana, K. J., van Ginneken, B., Nayar, S. K., and Koenderink, J. J. (1999b). Reflectance and texture of real-world surfaces. *ACM Trans. Graph.*, 18(1):1–34.

Day, E. A., Taplin, L. A., and Berns, R. S. (2004). Colorimetric characterization of a computer-controlled liquid crystal display. *Color Research and Application*, 29:365–373.

Dong, Y., Tong, X., Pellacini, F., and Guo, B. (2012). Printing spatially-varying reflectance for reproducing hdr images. *ACM Trans. Graph.*, 31(4):40:1–40:7.

Ferwerda, J. A., Pellacini, F., and Greenberg, D. P. (2001). Psychophysically based model of surface gloss perception. In *Photonics West 2001-Electronic Imaging*, pages 291–301. International Society for Optics and Photonics.

Fleming, R. W., Dror, R. O., and Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, 3(5).

Foo, S. C. (1997). A gonioreflectometer for measuring the bidirectional reflectance of material for use in illumination computation. Master's thesis, Cornell University.

García, P. A., Huertas, R., Melgosa, M., and Cui, G. (2007). Measurement of the relationship between perceived and computed color differences. *J. Opt. Soc. Am. A*, 24(7):1823–1829.

Hunter and Harold (1987). *Measurement of Appearance*. Chapter 6 (Gloss), 2nd edition.

Lafortune, E. P. F., Foo, S.-C., Torrance, K. E., and Greenberg, D. P. (1997). Non-linear approximation of reflectance functions. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 117–126, New York, NY, USA. SIGGRAPH '97, ACM.

Löw, J., Kronander, J., Ynnerman, A., and Unger, J. (2012). Brdf models for accurate and efficient rendering of glossy surfaces. *ACM Trans. Graph.*, 31(1):9:1–9:14.

Mantiuk, R., Krawczyk, G., Mantiuk, R., and Seidel, H.-P. (2007). High dynamic range imaging pipeline: Perception-motivated representation of visual content. In *Human Vision and Electronic Imaging XII*, volume 6492, San Jose, USA. SPIE.

Marschner, S. R., Westin, S. H., Lafortune, E. P. F., Torrance, K. E., and Greenberg, D. P. (1999). Image-based brdf measurement including human skin. In *Proceedings of the 10th Eurographics conference on Rendering*, EGWR'99, pages 131–144. Eurographics Association.

Matusik, W. (2003). *A Data-Driven Reflectance Model*. PhD thesis, Massachusetts Institute of Technology.

Matusik, W., Ajdin, B., Gu, J., Lawrence, J., Lensch, H. P. A., Pellacini, F., and Rusinkiewicz, S. (2009). Printing spatially-varying reflectance. *ACM Trans. Graph.*, 28:128:1–128:9.

Matusik, W., Pfister, H., Brand, M., and McMillan, L. (2003). A data-driven reflectance model. *ACM Transactions on Graphics*, 22(3):759–769.

Montag, E. D. (2006). Empirical formula for creating error bars for the method of paired comparison. *Journal of Electronic Imaging*, 15(1):010502.

Ngan, A., Durand, F., and Matusik, W. (2005). Experimental analysis of BRDF models. In *Proceedings of the Eurographics Symposium on Rendering*, pages 117–226. Eurographics Association.

Ngan, A., Durand, F., and Matusik, W. (2006). Image-driven navigation of analytical BRDF models. In *Symposium on Rendering*, pages 399–407, Nicosia, Cyprus. Eurographics Association.

Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., and Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance. Technical report, NBS Monograph 160.

Obein, G., Bousquet, R., and Nadal, M. E. (2005). New nist reference goniospectrometer. In *Proc. SPIE*, volume 5880, pages 58800T–58800T–10.

Obein, G., Knoblauch, K., and Vieot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision*, 4(9).

Oren, M. and Nayar, S. K. (1994). Generalization of lambert's reflectance model. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, SIGGRAPH '94, pages 239–246, New York, NY, USA. ACM.

Pellacini, F., Ferwerda, J. A., and Greenberg, D. P. (2000). Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 55–64. SIGGRAPH '00, ACM.

Pharr, M. and Humphreys, G. (2010). *Physically Based Rendering, Second Edition: From Theory To Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition.

Phillips, J. B., Ferwerda, J. A., and Luka, S. (2009). Effects of image dynamic range on apparent surface gloss. In *17th Color Imaging Conference*, pages 193–197.

Phong, B. T. (1975). Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317.

Reinhard, E., Stark, M., Shirley, P., and Ferwerda, J. (2002). Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 21(3):267–276.

Ren, P., Wang, J., Snyder, J., Tong, X., and Guo, B. (2011). Pocket reflectometry. In *ACM SIGGRAPH 2011 papers*, SIGGRAPH '11, pages 45:1–45:10, New York, NY, USA. ACM.

Schlick, C. (1994). An inexpensive brdf model for physically-based rendering. *Computer Graphics Forum*, 13:233–246.

Stark, M. M., Arvo, J., and Smits, B. (2005). Barycentric parameterizations for isotropic brdfs. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):126–138.

Vangorp, P. and Dutré, P. (2008). Shape-dependent gloss correction. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pages 123–130, New York, NY, USA. ACM.

Vangorp, P., Laurijssen, J., and Dutre, P. (2007). The influence of shape on the perception of material reflectance. *ACM Trans. Graph.*, 26.

Wang, C.-P., Snavely, N., and Marschner, S. (2011). Estimating dual-scale properties of glossy surfaces from step-edge lighting. In *Proceedings of the 2011 SIGGRAPH Asia Conference*, SA '11, pages 172:1–172:12, New York, NY, USA. ACM.

Ward, G. and Eydelberg-Vileshin, E. (2002). Picture perfect RGB rendering using spectral prefiltering and sharp color primaries. In *Proceedings of the 13th Eurographics workshop on Rendering*, EGRW '02, pages 117–124, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.

Ward, G. J. (1992). Measuring and modeling anisotropic reflection. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 265–272, New York, NY, USA. SIGGRAPH '92, ACM.

Westlund, H. B. and Meyer, G. W. (2001). Applying appearance standards to light reflection models. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '01, pages 501–51., New York, NY, USA. ACM.

Weyrich, T., Peers, P., Matusik, W., and Rusinkiewicz, S. (2009). Fabricating microgeometry for custom surface reflectance. In *ACM SIGGRAPH 2009 Papers*, SIGGRAPH '09, pages 32:1–32:6, New York, NY, USA. ACM.

Wills, J., Agarwal, S., Kriegman, D., and Belongie, S. (2009). Toward a perceptual space for gloss. *ACM Trans. Graph.*, 28(4):103:1–103:15.